Response to interactive comment by Anonymous Referee 1 on "Estimates of CO2 fluxes over the City of Cape Town, South Africa, through Bayesian inverse modelling" by Alecia Nickless et al.

Alecia Nickless, Peter J. Rayner, Francois Engelbrecht, Ernst-Günther Brunke, Birgit Erni, and Robert J. Scholes

We would like to thank the referee for their time and thorough consideration of the paper. We would like to thank the referee for their support of this work. With regards to the issues raised concerning the flow of the manuscript and redundancies highlighted by the referee, we have taken these comments on board, and performed a thorough rewriting of the manuscript which presents a more cohesive and focused presentation of the results of this reference inversion. The main purpose of this paper is to introduce the inversion framework used for the Cape Town inversion and to present the initial results for the reference inversion. The manuscript was rewritten to reinforce this focus. The sensitivity analyses will be presented in a related paper, and will rely on this introduction of the methodology. This paper focuses on the specification of the covariance matrices. Originally we had intended to include this work together with the reference inversion, but it soon became apparent that it would be too much content. Some of the discontinuities in the original manuscript are likely due to this change in focus of the paper. This was corrected in the revised manuscript.

As the manuscript has largely been rewritten to improve clarity, we will not focus on individual sentences identified by the referee, as many of these will be entirely changed or dropped from the manuscript.

Firstly we will address the main scientific concerns (labelled A to D by the referee).

The referee states *"A.)Figures 9a,c and 10a,c indicate very large misfits between the model and the measurements when using the prior fluxes while, as highlighted by the manuscript itself, one does not expect a very large variability of the CO2 in the area. How is it possible to get such an amount of misfits larger than 30 ppm (and in the range 100 to 200 ppm) while the data nearly never reach an excess of 30 ppm over their baseline, and while the measurement stations seem to be quite distant from the city major point sources? The manuscript ignores that such prior misfits strongly question the reliability of the atmospheric transport modeling framework, and thus of the inversion system.*

The authors say that the model "shows ability to track local events at the sites" but it is impossible to assess on Figures 9 and 10. Furthermore, given the very large size of the control vector (and thus the very large number of degrees of freedom in the inversion system), it is not really surprising to see that the inversion manages to fit the data to a far better extent. I find it difficult to take it as a demonstration that the atmospheric transport model is reliable. In particular, opposed to what is said on page 1 of the supplementary material ("some evidence to provide confidence in the modelled meteorology is provided in this section"), the figures 1 and 2 of this supplementary material strongly question this reliability, displaying very large misfits between the modeled and measured timeseries of wind speed, with a weak correlation between them. The difference between

the height of the measurement sites and the model vertical representativeness can hardly explain such misfits and even less such a low correlation. This requires a deeper analysis or better insights regarding the skill of the 1 km resolution CCAM model at city scale from studies like Engelbrecht et al. (2009, 2011).

By gathering the night-time and daytime data both in their analysis and in the inversion system, the authors do not help investigating this issue. We can assume that the largest misfits are obtained at night. However, opposed to what is said on page 1 line 12 ("Night-time observations were included, but allocated much larger errors compared to the daytime observations"), figure 5 indicates that the increase of the observation error at night is likely far from sufficient to cope with the increase of model errors at night. I think that the authors have overestimated their ability to assimilate night-time data. Analysis of the misfits between the modeled and measured CO2 and of the corrections to the fluxes applied by the inversion at night vs day could help investigate this topic."

The issues raised here are concerning the poor match between the prior modelled concentrations and the observed concentrations, as well as the results of the attempt at validating the wind information from CCAM in the supplementary material. Lastly the justification of the allocated uncertainties for the night-time observations is questioned.

The referee is concerned with the reliability of the atmospheric transport modelling framework. In retrospect, presenting the adhoc validation of the CCAM wind fields was unwise. Unfortunately the wind measurements we have available to us were not adequate to compare with the modelled winds from CCAM. The locations of the weather stations were far from the sites, except for Cape Point. All the locations for the weather station sites would not have been representative of the 1 km by 1 km pixel from which the modelled winds would have been extracted, or were situated in locations that would have been strongly influenced by the local topography or built environment. The anemometer located at the top of the Cape Point measurement tower is subject to winds influenced by the surrounding topography (Figure 1). As stated in Whittlestone et al. (2009), the single point measurements at Cape Point were strongly influenced by local topography and measured wind directions showed little correlation with the true source of the air mass: "Two superficially attractive selection criteria proved to be ineffective. One, wind direction, was so perturbed by local topography that there was no correlation of the measured wind direction with the bearing of the source of more distant trace gases from the critical north to east sector. The other, back trajectories, were effective in determining if contact with the southern African continent had occurred, but the indicated time and location of land contact was highly inaccurate. We make use of the Cape Point site for background concentration information only, and the inversion does not rely on correctly modelling transport to this particular site. Our two sites are slightly less extreme in the surrounding topography, but nonetheless, still pose significant challenges for obtaining wind speed and direction measurements that would be representative of a grid square produced by a regional climate model. All the weather station sites are located near the shoreline, therefore subject to sea-breeze variations.

Figure 1. Image of topography around Cape Point GAW station. http://www.imk-ifu.kit.edu/319.php

To justify the use of CCAM to provide modelled winds, we rely on previous studies which have used this model for atmospheric transport modelling in our target area (Whittlestone et al., 2009), and studies which have validated CCAM at different spatial resolutions (Engelbrecht et al., 2009; Roux , 2009; Engelbrecht et al., 2011, 2013, 2015). In particular, CCAM has been able to satisfactorily recreate present-day rainfall totals and the rainfall seasonal cycle, as well as circulation patterns over South Africa (Engelbrecht et al., 2009), and has been able to simulate with some success mid-tropospheric closed-lows and extreme rainfall events (Engelbrecht et al., 2015). CCAM has been validated over the Stellenbosch wine producing area, which falls within the domain of this inversion, with respect to temperature, relative humidity and wind speed at six different stations within this region (Roux , 2009). Validating the wind product from CCAM in a rigorous manner is beyond the scope of this paper. Our interest is in the impact of the estimates of the uncertainties on the inversion results, and we have focussed our sensitivity analyses on this question. The discussion in the manuscript reviewing CCAM's capabilities has been expanded to give a more thorough examination of this literature (section 2.4 in the revised manuscript).

Whittlestone et al. (2009) used CCAM to generate the wind data as well as to perform the transport modelling over Cape Point for their investigation. We instead relied only on the wind and other climate variables from CCAM, such as temperature which is well validated (Engelbrecht et al., 2009), and used the well known Lagrangian Particle Dispersion Model (LPDM) (Uliasz, 1994; Lauvaux et al., 2009; Lauvaux et al., 2016) for the transport modelling. This model has been used for several inversion exercises at various spatial scales.

The more likely candidate for the poor agreement between the observed and prior modelled concentrations is the specification of the prior fossil fuel and net ecosystem exchange (NEE) fluxes. The information available to perform a spatially and temporally disaggregated fossil fuel emission inventory for the City of Cape Town is limited. In addition, Cape Town is a city with stark inequalities between the population subgroups. For cities in developed countries it can generally be assumed that almost all heating and lighting will be generated through electricity consumption, but for many of the communities in Cape Town, this is not the case. These communities will rely heavily on raw fossil fuel burning for heating and lighting. Accessing this information and converting this into emissions based on the assumed behaviour of people is not factored into standard inventory analyses and is beyond the scope of this paper. Therefore, the residential emissions are a large contributor to the fossil fuel emission budget as well as one of the largest contributors to the uncertainties in the fossil fuel flux. The relevant question is whether we have successfully captured these uncertainties in specification of the prior uncertainty. The discussion of this issue has been included in section 2.9.1.

Another significant uncertainty is the simulation of the Fynbos biome. This biome is biodiverse, with many endemic species, but covers a relatively small area in South Africa, but a significant area within the domain of the inversion. The fynbos biome is poorly represented by dynamic vegetation models. The land atmosphere exchange model CABLE was selected to couple with CCAM due to its use and development in regions of Australia which share similar characteristics to the savanna biome in South Africa, which has a coverage of over 50%. Its ability to simulate respiration and photosynthesis in the fynbos region is largely untested. This discussion has been expanded in the revised manuscript in sections 2.7 and 2.9.2.

We can test whether our assumed uncertainties are consistent with the prior misfit in concentration (see Michalak et al. (2005) for details). The magnitude of the discrepancies that we obtained between the observed concentrations and the modelled concentrations are expected. If we calculate the matrix $\mathbf{H}\mathbf{C}_{\mathbf{s_0}}\mathbf{H}^T + \mathbf{C}_{\mathbf{c}}$, and assume $\mathbf{H}$ is correct, the uncertainty in prior fluxes is projected into uncertainty in modelled concentrations. The square root of the diagonal elements have a similar distribution to the absolute mismatches between the observed and prior modelled concentrations. This has been included at the beginning of section 3.1.

To confirm that the prior information is the major cause of the disagreement (although not having perfect transport modelling is to blame as well), we performed a sensitivity analysis where the NEE estimates were averaged over the domain, and the prior estimates for NEE set as the overall average for all pixels. Similarly, the uncertainty of the NEE estimates was set as the overall average of the net primary productivity (NPP) estimates. At the Hangklip site, which is dominated by the NEE contributions, this led to far smaller discrepancies between modelled and observed concentrations (compare Figures 2 to 3). Therefore it appears that CABLE may be overestimating the amount of photosynthesis or respiration (or both) that are taking place in the region. On the other hand, Robben Island, which is far more influenced by the Cape Town fossil fuel emissions, did

not show any improvement in the discrepancies between the prior and observed concentrations (compare Figures 4 to 5). In the sensitivity analysis paper, we present further tests based on the specification of the fossil fuel emissions and NEE fluxes and their uncertainties. These timeseries plots have been included in revised manuscript (Figures 6 and 7), as well as in the supplementary material in section 1.4.

The plots of the residuals between the modelled and observed concentrations are not very different between day and night (see the Figures 2 and 4). Therefore it does not appear that the difficulty in modelling transport at night is leading to significantly larger discrepancies between the observed and modelled concentrations, as suggested by the referee. Specifying uncertainties that are too small for the night time model errors does not appear to be leading to the large discrepancies we are observing. The time series plots have been given for only a short period in order to avoid squashing the time axis, and to better demonstrate the inversions ability to obtain posterior modelled concentrations that are much closer to the observations. Although this is entirely expected, we wanted to demonstrate that as far as optimising the flux estimates to better match the observed concentrations, the inversion was at least achieving this. The expanded time series plots better show the ability of the inversion to track local "pollution" events.

The revised manuscript now includes more discussion on the discrepancy between the observed and prior modelled concentrations. To deal with the problems related to viewing the observed and modelled concentration data in Figures 9a,c and 10a,c, the time series plots have been altered. They are expanded over several panels in order to stretch out the time axis, allowing the reader to better assess the discrepancies and similarities between the observed and modelled concentrations, as shown in Figures 2 to 5. We included the diurnal plots to demonstrate how the prior and posterior modelled concentrations were differing from the observed at different hours of the day.

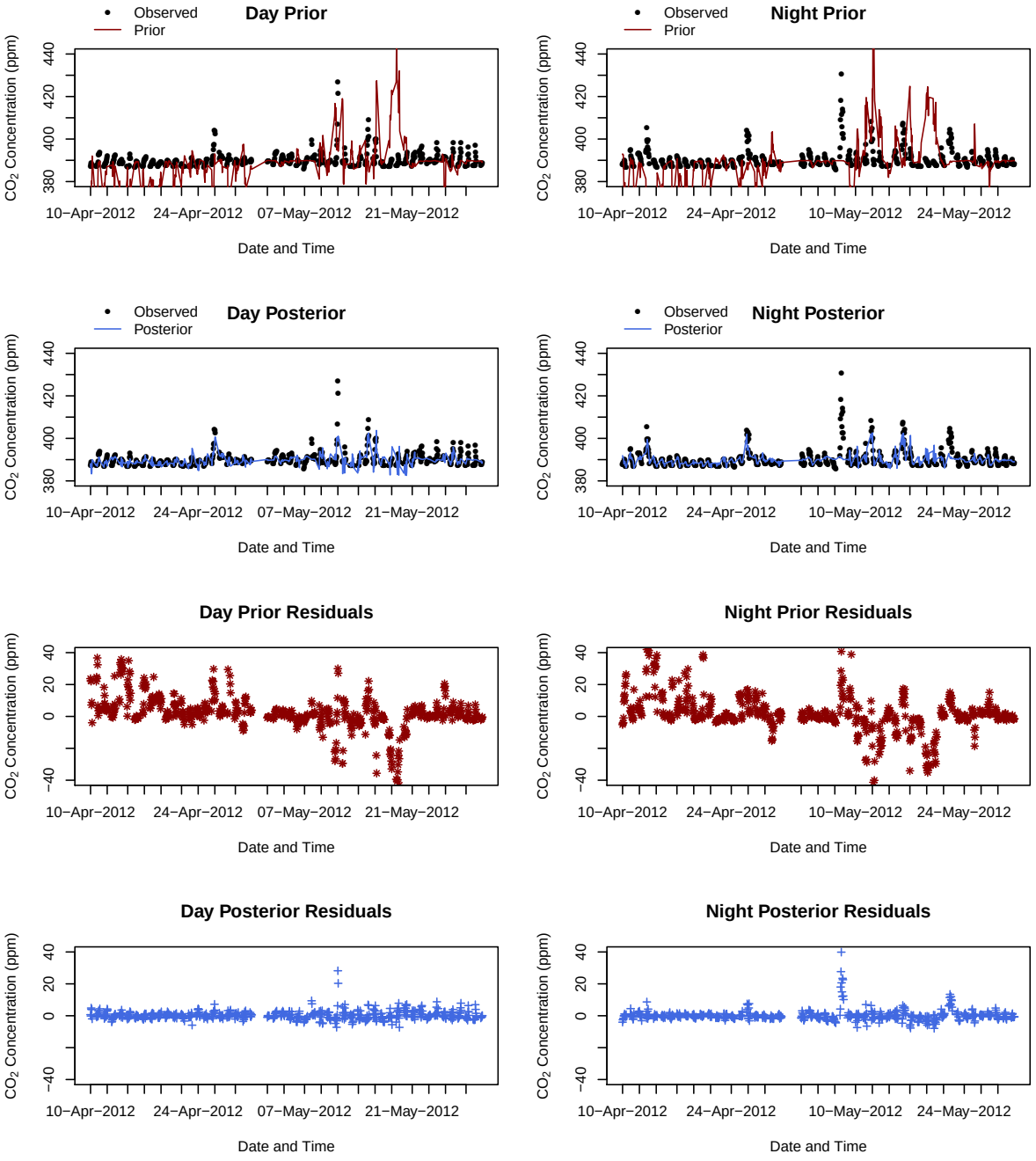


Figure 2. A time series of the observed and modelled CO_2 concentrations for April and May 2012 at Hangklip separated by day and night

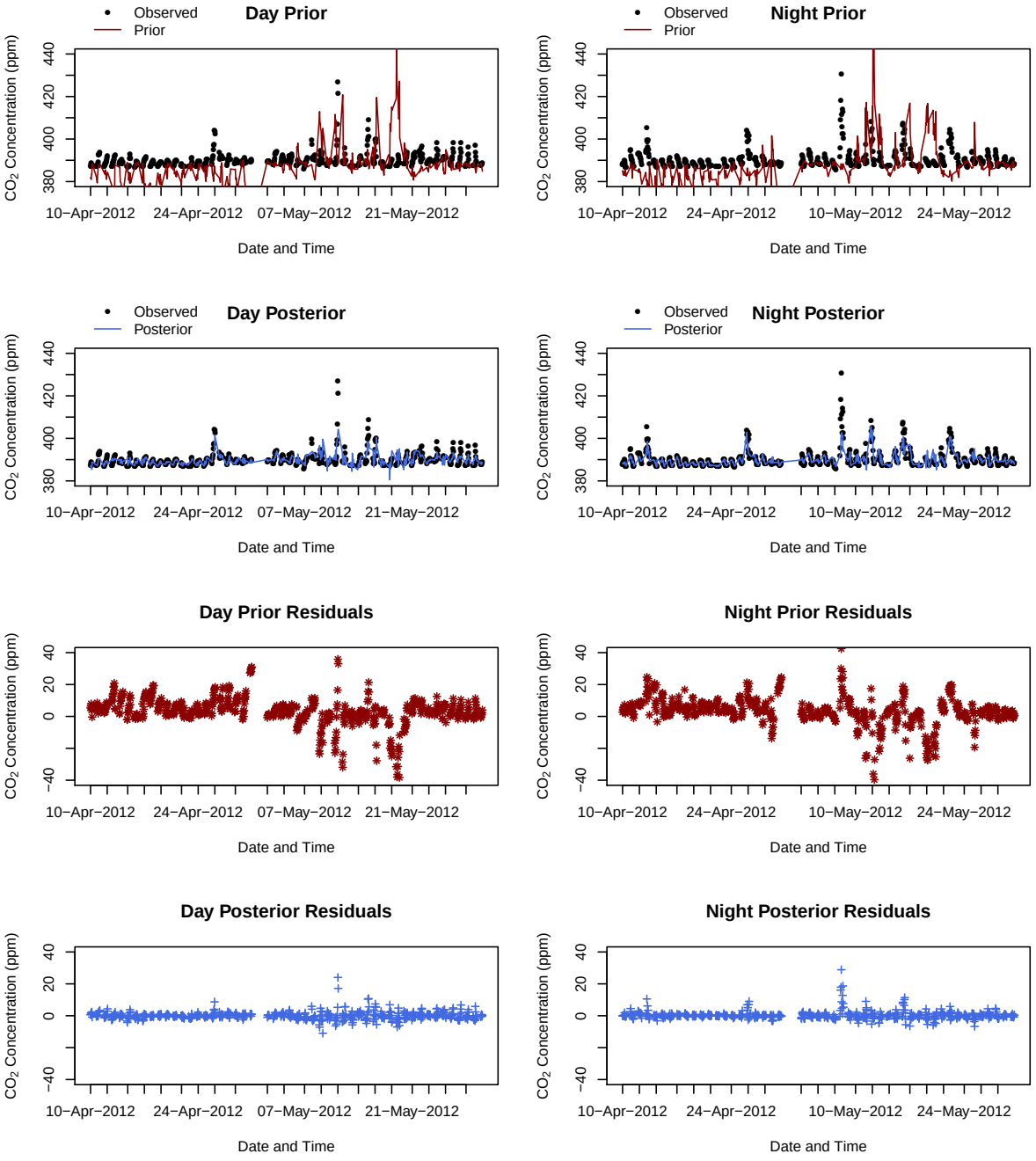


Figure 3. A time series of the observed and modelled CO_2 concentrations for April and May 2012, where the prior estimates of NEE have been set as the average NEE over all pixels, at Hangklip separated by day and night

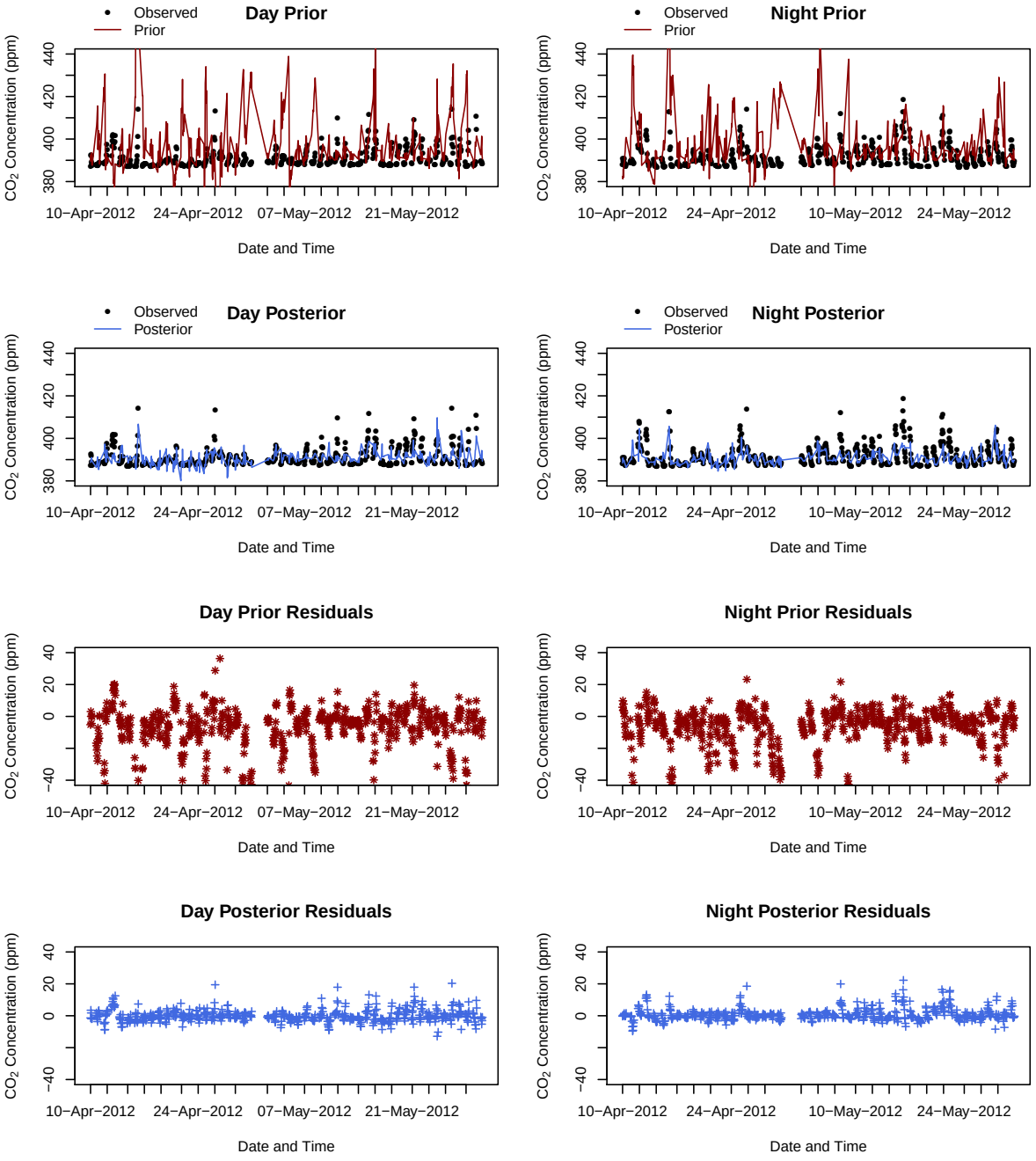


Figure 4. A time series of the observed and modelled CO_2 concentrations for April and May 2012 at Robben Island separated by day and night

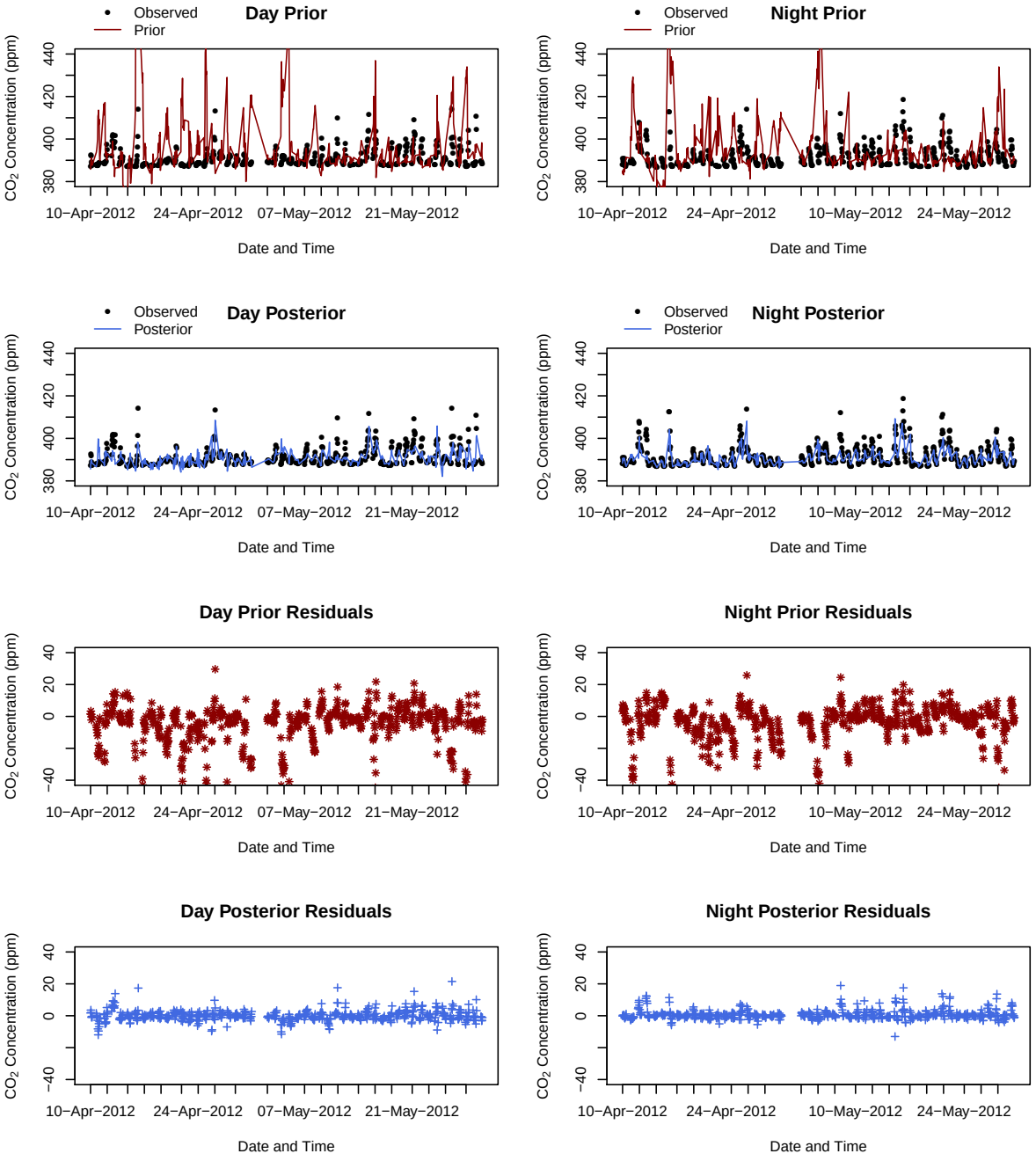


Figure 5. A time series of the observed and modelled CO_2 concentrations for April and May 2012, where the prior estimates of NEE have been set as the average NEE over all pixels, at Robben Island separated by day and night

The second major issue that the referee raised was:

"B) One of the main discussion of the manuscript is related to the lack of distinction between the anthropogenic and natural fluxes in the inversion results. What puzzles me is that the analysis of the posterior error covariance matrices should be a very helpful tool to feed such a discussion. The authors display correlations between uncertainties in the emissions at a given pixel and the NEE field in Figure 16 but ignore them when conducting this discussion.

Actually, Figure 16 shows correlations that are very close to 0, which undermines the assumption of the lack of distinction. When looking at the station locations and at the situation of the city vs. the areas of high NEE (which are also the areas of high prior uncertainties in the NEE and of high corrections to the NEE in the inversion), it is difficult to understand why the separation between the NEE and the anthropogenic emissions should be so problematic for the inversion. The authors have used very large prior uncertainties in the NEE, and the NEE dominates the mean diurnal cycles of the stations. This explains why, on the first order, the inversion focuses on the NEE rather than on the anthropogenic emissions (from that point of view, it would not be a problem of distinction between NEE and anthropogenic emissions, but rather a problem of detection of the emissions despite the dominating signal from NEE). But, according to figures 3 and 4, this should not prevent the inversion from getting a signal that is dominated by the anthropogenic emissions when the wind blows roughly from one station to the other one through Cape town. Paradoxically, the type of "gradient approach" that the author assume to be useless for their study case (p13 line18 p61 lines7-9) may help them to cope with the NEE signal. All of this needs to be better analyzed. The analysis of the variations of the modeled contribution from the NEE vs that from Cape Town at the measurement sites could be very useful."

The referee is concerned about the conclusions regarding the difficulty of separating NEE and fossil fuel fluxes correctly in the inversion. This is an important concern although we do not agree with the metric the referee chooses. We believe that the correct metric is the uncertainty of the difference between the two flux components, not the correlation. To see this, imagine the case where two flux components from the same pixel are given independent priors (no prior correlation) and are not all visible to the observation network. Nothing will have been learned about the fluxes, including their difference, but the posterior correlation will be zero. Unfortunately, this is close to our present case for many pixels. There is not enough information available to the inversion for it to correctly assign contributions from NEE and fossil fuel fluxes. We solve for these contributions separately in the inversion so the flexibility exists for adjustments to be made to both of these fluxes while attempting to obtain sensible adjustments to the overall flux within the pixel. As the referee also mentions, this is made more difficult by the large uncertainty that we have assigned to the NEE estimates (more on this in the third issue raised).

The covariance between fossil fuel and NEE flux uncertainties are small because the uncertainties in the prior modelled concentrations that are attributed to the flux contributions ($\mathbf{HC_{s_0}H}^T$) are small relative to the uncertainties specified for the modelled concentration errors ($\mathbf{C_c}$). This is not because our prior uncertainty is small but because the transport Jacobian only projects fluxes from individual pixels weakly into modelled concentrations. As the uncertainty in the modelled concentration errors is decreased, the size of the posterior off-diagonal covariance elements between the fossil fuel and NEE flux uncertainties from the same pixel increases. This can easily be confirmed through the use of a toy inversion system using typical values for $\mathbf{H}$, $\mathbf{C_{s_0}}$ and $\mathbf{C_c}$ from our inversion system. The posterior variance of any linear combination of terms from the source vector

of the fluxes (including the difference between two fluxes from the same pixel) will always be reduced or (at worst) left unchanged relative to the prior variance of the same linear combination of elements (Jackson , 1979; Jackson and Matsu'ura , 1985). Observations may very well introduce correlations between flux components but this does not mean a reduction in the ability to distinguish between them.

If we define the distinction between the fossil fuel flux and NEE flux within the same pixel *i* as the variance of the difference between the fossil fuel and NEE fluxes $s_{f,i} - s_{NEE,i}$, this will be equal to the sum of the variances of these two fluxes minus twice the covariance between them: $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} - 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ will be negative. Therefore the posterior uncertainty of the difference in these fluxes is always going to be larger than the sum of the individual posterior flux uncertainties, but smaller than the prior uncertainty of this linear combination of terms. Therefore, although the off-diagonal terms may be small, the ability of the inversion under the current framework to detect between NEE and fossil fuel fluxes is limited as the posterior uncertainties are still large, and therefore the uncertainty of $s_{f,i} - s_{NEE,i}$ is large. If the covariance terms are small because, relative to the errors in the modelled concentrations the contribution of the uncertainty in the fluxes to the uncertainty in the modelled concentration is small, then the variance of $s_{f,i} - s_{NEE,i}$ is still going to be large due to the dominance of uncertainty in $s_{NEE,i}$.

On the other hand, when we aggregate these fluxes from the same pixel to get the total flux within a pixel $s_{f,i} + s_{NEE,i}$, the uncertainty of this term is equal to $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} + 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ is negative. We already know that the sum of any linear combination of sources will have a smaller uncertainty after the inversion, but when we aggregate fluxes from the same pixel, the uncertainty of this total is smaller due to the smaller posterior variances and also because the covariances are negative. This demonstrates that the value of the inversion is to reduce the uncertainty on each of the individual fluxes and to additionally reduce the uncertainty of the aggregation of the NEE and fossil fuel flux within the same pixel. The reduction in the uncertainty of the sum of fluxes within the same pixel is going to depend on the size of the uncertainty of the NEE flux, which is usually the larger uncertainty. To improve the ability of the inversion to estimate the total flux within a pixel, we need to improve the skill of the atmospheric transport and we need to reduce the uncertainty in the estimates of the NEE. As it stands, with a large prior uncertainty in the estimation of the NEE fluxes from the CABLE model which remains a large posterior estimate after the inversion, the distinction between the fossil fuel and NEE flux from the same pixel is not very different from the prior estimate of the variance of the difference between the fossil fuel and NEE flux, $C_{s_0(f,i;f,i)} + C_{s_0(NEE,i;NEE,i)}$.

We would like to thank the referee for the suggestion of adding the investigation of what contributions NEE and fossil fuel fluxes make to the modelled concentrations at each site. This analysis has now been included into the manuscript. This discussion is at the end of section 3.2. Figures 9 and 10 have now been included which provide information on the corrections that inversion is making to the contribution to the modelled concentration from the fossil fuel and NEE fluxes. An addition discussion is provided in the supplementary material (section 1.8).

We still have a long way to go before we can reliably estimate the fossil fuel fluxes from the City of Cape Town using this inversion framework, but our intention is to provide the building blocks for such an inversion system which would allow better products for fossil fuel emissions and NEE fluxes to be slotted in as they become available, while we work towards a reduction

in the misfit between the prior and posterior modelled concentrations. For example, we could replace the bespoke inventory analysis used here for the global 1×1 km ODIAC fossil fuel product (Oda et al. , 2017) which has recently been used for Indianapolis. This has been performed for the sensitivity analysis paper.

The referee mentions that the gradient approach may have been more useful in the Cape Town inversion than implied in the original manuscript. The main concern we had about using the gradient method in these circumstances was that the direction of travel of an air parcel between the two sites would not necessarily be in a straight line due to the topography of the Cape Town region. Therefore selecting observation pairs based on the wind direction would not necessarily have given the true gradient in concentration between the sites. We supplied a map of the sensitivities of the observations sites to the surface fluxes to show this in the original manuscript (now Figure 17 in the revised manuscript). To further add to this argument, we provided figures of the average wind speed and direction for the domain for each month in the supplementary material (section 1.3). These figures show that, in general, the wind direction would not be in our favour, and with only two sites, that would leave very little information to constraint the surface fluxes. When the wind is blowing from the Hangklip site, it curves northwards towards the interior and away from Cape Town. When the Robben Island site is observing clean marine air before coming into the Cape Town area from the Atlantic side, such as June 2013, the wind changes from the north westerly direction once it passes over Cape Town to a more northerly direction, missing the Hangklip site. These wind plots have replaced the original attempt at validating the CCAM wind data.

The third major issued related to the NEE uncertainty estimates. These were set to be large (relative to the NEE estimate), and we used the estimates of the productivity associated with the NEE estimate as the uncertainty. More precisely, which is made clearer in the manuscript, as the uncertainties were scaled so that the χ-squared goodness-of-fit statistic was closer to 1, the productivity values (NPP) were used to define the relative uncertainties between the NEE estimates. There is a typo in the original manuscript which states we used Net Ecosystem Productivity for the uncertainties where it should be Net Primary Productivity. The NPP fluxes were squared to give initial estimates of the variances. The fossil fuel and NEE variance estimates were then doubled so that the Chi-squared goodness-of-fit statistic was closer to 1, and these scaled variances provided the final uncertainties assigned to the prior fluxes. These scaled uncertainties were correctly displayed in prior information figures, but not made clear in the text.

"C) Regarding the prior uncertainties in the NEE, the relative values discussed in section 2.9.2 can be very large. They deserve some justification based on the CABLE validation studies, especially since they will be amplified by the multiplication of the prior uncertainties by a factor of 2 in section 2.11 (Figure 4 being misleading). I understand that when aggregating them over the modeling domain, we get much smaller relative values due to using a small spatial correlation length scale. However, could not it be an issue for the results at the pixel scale and, implicitly, for the control of the highly localized anthropogenic emissions (see the strange correction patterns in figure 13 and 14) ?"

We certainly agree that the large uncertainty we have in the NEE estimates produced by CABLE is a limitation and inhibits the ability of the inversion to apply corrections to the fossil fuel fluxes, in which we are the most interested. Modelling NEE in this region is uncertain. Wang et al. (2011) have shown that unless CABLE is closely calibrated for a specific system, it can lead to significant errors in the estimation of NEE. There are many land types in the region, which include the endemic fynbos biota, different types of agriculture, as well as fallow land. Although a great deal of work is being carried out to validate CABLE over the savanna biome, which covers much of South Africa, the information available for how well CABLE behaves for fynbos is limited. In the sensitivity analysis we will consider other models of NEE for this region, but for the reference inversion, we think that a conservative estimate of the error is best. Sensitivity analyses will consider the impact of reducing the uncertainty estimates to a smaller fraction of the NPP estimates from CABLE.

The final major scientific issue that the referee raises concerns the inventory analysis:

"D) I am not sure to understand the distribution of the emissions from Cape Town according to the author's inventory. p11 says "But of the carbon emissions due to energy usage, only 27% were attributed to the transport sector as a result of the carbon intensive usage of coal for electricity generation to provide almost all of the energy to the residential and commercial sectors in South Africa, which emit approximately 29% and 28%, respectively, of the total carbon emissions of CT (City of Cape Town, 2011)." Paying much attention to the terms "electricity generation" and "almost all of the energy", my understanding of this sentence is that there is a large number of coal power plants within the city bounds (otherwise the part of the emissions within the city due to the transport would be very high), which represent almost 60% of the city CO2 emissions, while the direct emissions from the residential and commercial areas should be very low. However, this seems strongly at odd with the figures

and discussions of this manuscript and this would be highly problematic for the atmospheric inversion. Could the authors clarify this point "

The referee is concerned about contributions to the fossil fuel budget from the different sectors of the city. The percentages that are quoted here come from a report on the energy consumption of the city, and do not entirely relate to the direct fossil fuel emissions from the city. Emissions from coal are small because most of the power generation capacity through coal occurs in the north eastern provinces of South Africa. Residential emissions are not negligible in Cape Town because many of the communities still rely on burning raw fossil fuels for heating, cooking and lighting. This discussion has been made clearer in the revised manuscript. Instead of discussing these energy statistics which are already discussed in the inventory paper, the discussion in the revised manuscript now reflects the percentage contributions as reflected in the inventory data available for this inversion. Using these statistics, the emissions from industry (based on the available fuel usage data) are 12.0%, 34.6% from vehicle road transport, 51.0% from the residential sector, and 2.4% from the airport and harbour. This discussion has been updated in section 2.6.

The referee lists a number of additional issues which have been corrected in the revised manuscript.

"*Section 2.1 makes a rough account of the traditional theoretical framework of the inversions (e.g. sentences like "If we assume a Gaussian error distribution for the surface fluxes and concentrations we obtain the following cost function for our least squares problem" on page 6). Throughout the manuscript, the covariances between uncertainties in fluxes are often called covariances between fluxes.*"

The description of the Bayesian inversion framework has been rewritten in the revised manuscript (section 2.1). With regards to the covariances, the wording has been clarified.

"*- in the abstract and first sections, the text introduces the concept of "boundary concentrations" without specifying that the boundaries relate to the modeling framework. This becomes problematic when explicitly speaking about the sensitivities of the measurements to the boundary concentrations (e.g. l23-24 p5)*"

We have made this description clearer in the manuscript to reflect that we are referring to the concentrations at the boundaries of the modelling domain when we refer to "boundary concentrations".

"*- the text often uses the terms "sources" and "emissions" while speaking about (natural) fluxes that can be negative*"

This has been corrected in the manuscript.

"*- p10 line17-19 the sensitivity is not the influence, H is not HTtranspose*"

This has been corrected and the description of the methodology made clearer".

"*- p11 "and allowed for small scale transport features to be maintained in H": using a coarse resolution control vector does not remove the small scale transport features in H*"

This has been removed from the sentence.

""- Section 2.11 states that the error covariance matrix that is underestimated in the first configuration according to the chi test is necessarily B ("and values greater than one indicate that the variance prescribed is lower than it should be and therefore the posterior estimates will be over-constrained by the prior fluxes") while it could be R (and actually some of the results favor the assumption that it is R)."

As discussed above in response to the referees concerns regarding the atmospheric transport modelling, we think that the prior information is the most uncertain in our inversion system. We consider other configurations of the covariance matrix for the errors in the modelled concentrations in the sensitivity analysis paper.

The individual sentences identified by the referee on pages C6 to C8, and similarly problematic sentences elsewhere in the manuscript, have been amended or cut from the revised version.

"- The design of the figures should be improved. The location and name of the sites are hardly visible in figure 1. Labels are too small and the fields are fuzzy (mainly due to the choice of the colorbars) in figures like Figure 3. The choice of the colors in figures like Figure 8 is poor: on my screen, it is really hard to distinguish between the different curves. In figures like Figure 9, it is impossible to analyze the different timeseries since they are compressed along the x axis (with nighttime and daytime data mixed together) and most of the measurements are hidden behind the model patches. In most of the figures, there is a lack of subtitle and legends to help the reader while the captions are sometimes quite complex (e.g. the legend of figure 15). Therefore, in general, the figures are very difficult to read. In Figure 15: it is difficult to see the pixel against which covariances are computed."

Figures in the manuscript have been replotted to address the concerns highlighted above. The time series have been expanded to multiple panels as explained early, and the day and night-time concentration and residuals separated. Better use has been made of legends and subtitles in the figures, to avoid including too much information in the caption. The maps have been improved so that labels are easier to read. The colour scheme has been amended to the rainbow colour scheme. In the original manuscript, the purpose of deliberately not using a smooth colour scheme was to allow very large fluxes, like those from point sources, to be distinct from smaller fluxes, while still allowing differences in these smaller fluxes to be distinguished, such as subtle differences in the residential and transport fossil fuel fluxes. We have now used the log scale where necessary, instead of customised graduations in the colour bar. The pixel against which covariances are computed in Figures 15 and 16 has been clearly identified (now just Figure 15 in the revised manuscript).

"- The notations used in several equations are not really optimal. At least, they do not help understand the meaning of the variables, e.g. Etrans in equation 11 which refers to a subcomponent of the transport error, while Eobs refers to another part of the transport error, and not to the observation error (which is the sum of the transport and measurement errors). Eq 8 is not really adapted to equations 1 and 6. Eq 9 and 11 are informal."

The notation for these equations has been amended to be more consistent with the rest of the manuscript. These are now equations 9, (10 and 11) and 13.

"- *There are too many significant digits in tables 3 and 4 which makes these tables difficult to read. Would not it be better to show the content of these tables using plots ? I do not understand why the authors produce distinct sections (3.2.3 and 3.2.4) and tables (3 and 4) for the variations of the 1-week mean and 1-month mean flux budgets. This is a source of redundancies and I don not think that they manage to bring specific insights for each of the two timescales.*"

Table 3 has been converted into a time series of box plots to show the prior and posterior distribution of pixel-level fluxes over the domain. These fluxes are also presented as maps in the supplementary material (section 1.7). Table 3 was produced to deliberately consider the summary statistics of the pixel-level flux estimates within the domain, and to show how these summary statistics differ between months. The objective of section 3.2.2 was to address how the inversion was updating fluxes at the pixel level, and from month to month. In this section we were comparing the prior estimates $\mathbf{s_0}$ with the posterior estimates $\mathbf{s}$. Therefore the unit of the fluxes are $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, which is unit of the fluxes solved for by the inversion. Sections 3.2.3 and 3.2.4 were aimed at addressing what impact the inversion had on the total flux of $\mathrm{CO_2}$ over the whole domain. These two sections have been merged. The purpose of the month estimates was originally for the sensitivity analysis to compare between configurations.

As discussed earlier, the inversion has the most impact in reducing uncertainty when aggregating over fluxes in the domain, and the main objective of Table 4 is to show this uncertainty reduction. When aggregating fossil fuel and NEE fluxes, and aggregating over all pixels, the reduction in the posterior variances and the negative posterior covariances brought about by the inversion can be taken advantage of to produce posterior total estimates which have associated uncertainties which are much smaller than those of the prior total estimates.

"- *The acronyms CT and CBD are not defined explicitly.*"

This has been corrected.

"- *Section 2.2. and Equation (6) are confusing regarding the composition of the control vector (regarding the fact that the inversion solves for the fluxes at the transport model spatial resolution and regarding the control of the average conditions for each of the 4 lateral boundaries). We need to guess it from the numerical derivation of the size of the control vector or wait for sections 2.5 to get clearer details. The situation is similar regarding the fact that 1-month inversions are conducted to cover the 13 month period.*"

This has been made clearer in the revised manuscript.

"- *there is a problem with the order of the citations (see Tarantola (2005) and Enting (2002) and Lauvaux et al., 2016; Bréon et al., 2015 on page 5)*"

This has been corrected.

"- the percentile filtering technique at Cape point and its impact on the timeseries at this site is not well detailed (e.g. on which time windows, at which timescales is it applied ?), while the station can be influenced by Cape Town, and by the NEE in the region covered by the modeling domain. This is perturbing since the system controls the North and East boundary conditions that are inland (and thus separated from the Cape point station by large areas of NEE and potentially influenced by even larger areas of NEE outside the modeling domain) and since it uses the data filtered at Cape point to provide a prior value with a low prior uncertainty to these conditions."

More discussion has been added on the percentile filtering technique (seciont 2.8 of the revised manuscript). A detailed description of this technique is provided in (Brunke et al., 2004): "The background signal at Cape Point, obtained from a percentile filtering technique (Brunke et al., 2004), was used as the prior estimate of the concentrations at each of the four domain boundaries. The percentile filtering technique removes data influenced by the continent or anthropogenic emissions. Two 11-day moving percentiles, which are adjustable by tuneable factors, control the upper and lower threshold limits. This results in a subset of background measurements from Cape Point represented by a narrow concentration band contained within these limits. This filter, when applied to the Cape Point CO_2 measurements, selects approximately 75% of the data. The percentile-filtering technique has been shown to compare well with the more robust method of using contemporaneous radon (^{222}Rn) measurements to differentiate between marine and continental air.".

The purpose of the percentile filtering technique is to remove those measurements which are strongly influenced by either anthropogenic emissions from Cape Town (which are observed very seldom by the tower) or those measurements strongly influenced by biospheric uptake of CO_2. This discussion has been expanded in the manuscript. We have deliberately made the margin of the domain around the City of Cape Town large, which allows the inversion system to solve for fluxes at large distances from the City, rather than relying on estimates of the concentrations at the boundary. Therefore when the air arriving at the measurement site originates from the north the inversion can account for uptake by correcting the far-field fluxes and leave the concentration at the far North or East boundaries relatively unchanged. In this way, the boundary CO_2 concentrations act more like baseline concentrations.

"- p20: the discussion on the representation error ignores the part of this error due to the difference of spatial representativeness between the measurements and the model ("We did not account any further for aggregation or representation errors as we did in the network design, as we were running the inversion at the same spatial scale as the transport model.")"

The representation error is accounted for in the 2 ppm and 4 ppm assigned to daytime and night-time concentrations respectively. The discussion in the manuscript has been altered to better reflect this. The representation error occurs due to errors in the transport modelling. The distinction between aggregation error and representation error has been corrected in the manuscript (section 2.10).

References

Brunke, E. -G, Labuschagne, C., Parker, B., Scheel, H. E, and Wittlestone, S.: Baseline air mass selection at Cape Point,South Africa: application of ^{222}Rn and other filter criteria to CO_2, Atmos. Environ., 38, 5693–5702, doi:10.1016/j.atmosenv.2004.04.024, 2004.

Engelbrecht, F. A., McGregor, J. L. and Engelbrecht, C. J.: Dynamics of the conformal-cubic atmospheric model projected climate-change signal over southern Africa, Int. J. Climatol., 29, 1013-1033., doi: 10/1002/joc.1742.29., 2009.

Engelbrecht, F. A., Landman, W. A., Engelbrecht, C. J., Landman, S., Bopane, M. M., Roux, B., McGregor, J. L., and Thatcher, M.: Multi-scale climate modelling over Southern Africa using variable-resolution global model, Water Research Commission 40-Year Celebration Conference, Kempton Park, 31 August - 1 September 2011, doi: 10.4314/wsa.v37i5.2, 2011.

Engelbrecht, C. J., Engelbrecht, F. A. and Dyson, L. L.: High-resolution model-projected changes in mid-tropospheric closed-lows and extreme rainfall events over southern Africa. Int. J. Climatol., 33, 173–187, doi: 10/1002/joc.3420, 2013

Engelbrecht, F., Adegoke, J., Bopape, M-J., Naidoo, M., Garland, R., Thatcher, M., McGregor, J., Katzfey, J., Werner, M., Ichoku, C. and Gatebe, C.: Projections of rapidly rising surface temperatures over Africa. Env. Res. Letters., 10(8), 085004, doi: 10.1088/1748-9326/10/8/085004, 2015.

Jackson, D. D.: The use of a priori data to resolve non-uniqueness in linear inversion, Geophys. J. R. astr. Soc., 57, 137–157, 1979.

Jackson, D. D and Matsu'ura, M.: A Bayesian approach to nonlinear inversion, J. Geophys. Res. 90 (B1), 581–591, 1985.

Landman, S., Engelbrecht, F. A, Dyson, L., Engelbrecht, C., J., and Landman, W., A.: A short-range ensemble prediction system for South Africa, Water SA, 38(5), 765–774, doi: 10.4314/wsa.v38i5.16, 2012.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, doi: 10.5194/bg-6-1089-2009, 2009.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337–354, doi: 10.5194/acp-12-337-2012, 2012.

Lauvaux, T., Miles, N. L., Deng, A. Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), J. Geophys. Res. Atmos., 121, 5213–5236, doi: 10.1002/2015JD024473, 2016.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 110, D24107, doi: 10.1029/2005JD005970, 2005.

Oda, T., Lauvaux, T., Lu, D., Rao, P., Miles, N. L., Richardson, S. J. and Gurney, K. R.: On the impact of granularity of space-based urban CO_2 emissions in urban atmospheric inversions: A case study for Indianapolis, IN, Elem Sci Anth, 5, 28, doi: 10.1525/elementa.146, 2017.

Roux, B.: Ultra high-resolution climate simulations over the Stellenbosch wine producing region using a variable-resolution model, MSc Thesis, Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa, 106 pp., 2009.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics , Philadelphia, 2005.

Uliasz, M.: Lagrangian particle modeling in mesoscale applications, in: Environmental Modelling II, Computational Mechanics Publications, Southampton, UK, 71–102, 1994.

Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., van Gorsel, E., and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, J. Geophys. Res., 116, G01034, doi:10.1029/2010JG001385, 2011.

Whittlestone, S., Kowalczyk, E., Brunke, E. G., and Labuschagne, C.: Source Regions for CO_2 at Cape Point Assessed by Modelling 222Rn and Meteorological Data, Technical Report for the South African Weather Service, Pretoria, South Africa, 2009.

Response to interactive comment by Anonymous Referee 2 on "Estimates of CO2 fluxes over the City of Cape Town, South Africa, through Bayesian inverse modelling" by Alecia Nickless et al.

Alecia Nickless, Peter J. Rayner, Francois Engelbrecht, Ernst-Günther Brunke, Birgit Erni, and Robert J. Scholes

We would to thank the referee for their consideration of the paper and for their support of this work.

Referee 2 states: "*Overall I think this work makes an important contribution and should therefore be of interest for readers of ACP. Atmospheric inversions over cities are challenging and I appreciate the amount of work that has gone into this study. However, I think the manuscript can still be improved. The results section for example contains (too) many details including figures and tables. This makes it really difficult for the reader to focus on the main findings. Sentences are sometimes overly long and therefore hard to understand. Figure quality could also be improved.*"

The manuscript has been largely rewritten to improve the focus of the paper and improve clarity. Figures have been improved to be clearer, use better colour schemes, and to provide more information through legends rather than overly long and complicated captions. Certain sections have been removed or merged. For example, the supplementary section on the validation of the wind product from CCAM has been removed as the sites were generally not equipped or suited for this purpose. Instead more information on the general wind patterns within the domain has been provided, which is of more use in this context. The section 3.2.3 and 3.2.4 have been merged.

The referee makes the following specific comments:

"*Abstract:*

- please consider rephrasing the sentence stretching from L6 to L11 on P1."

The abstract has been completely rewritten.

"*- information such as "interquartile range" should be removed from the abstract (P1: L15, L16, L17, L18)*"

This has been removed from the abstract.

"*Main text:*

P5, L9: "CABLE was dynamically coupled to CCAM" I don't know how many times I read this phrase in the manuscript. Please avoid repetitions."

As in response to the remarks by Referee 1, these sorts of redundancies have been removed through a thorough revision of the manuscript. The only details which are repeated in a few places relate to the structure of the control vector, as this is a distinctive feature of the inversion framework, and results in the dimensions of the matrices $\mathbf{C}_{\mathbf{s_o}}$ and $\mathbf{H}$.

"P6, L19: surface source grid "point" ?"

What was meant here was "surface pixel"

"P10, L5-L23: you talk about the generation of sensitivity matrix H, but you also mention sensitivity matrix T, which is confusing"

This was a typo. It has been corrected in the revised manuscript. There is no matrix T.

"P11, L14-L16: I don't understand this sentence. Please rephrase using shorter sentences."

This paragraph has been rewritten:

Previous: "Since the surface sources are expressed as fluxes of carbon, the contribution to the concentration at the measurement site is expressed in the amount of carbon seen at the measurement site from a particular source. In the case of the boundary sources (or contributions from outside of the domain) which are given as concentrations, their contributions to the concentration at the measurement site are expressed as a proportion of their concentration, dependent on their influence at the receptor site. Ziehn et al. (2014) shows that by calculating the Jacobian which provides the sensitivities of observed concentrations to boundary concentrations, the boundary contribution can then be written as:"

Revised: "The fluxes from the surface pixels are expressed in $\mathrm{kg\ CO_2 m^{-2}\,week^{-1}}$ and are transformed through $\mathbf{H}$ into contributions to the concentration at the measurement site in units of ppm. The inversion solves for the concentrations at the boundary of the domain. Ziehn et al. (2014) shows that the Jacobian which provides the sensitivities of observed concentrations to boundary concentrations can be calculated as:"

"P11, L23: "For the network design ..." ? It is confusing that you refer to a network design here."

This has been removed.

"Figs.3+4: This figures need improvement. Please label the sub-plots. The information is provided in the caption, but it would be much easier to have it in the figure directly. The used colour scheme is probably not that helpful either."

These plots have been revised to make better use of subplot titles and legends to reduce the amount of text in the captions, and the text has been made clearer. The colour scheme has been altered to improve clarity. The colour scheme in the original manuscript was deliberately chosen so that 1.) positive, negative and zero fluxes would be distinct; 2.) and the "jagged" graduation in colours so that the very large fossil fuel fluxes at point sources could be distinct, but also the smaller fluxes, such as those for transport and residential sources. These are smaller, but more numerous than the point sources, and we wanted these to be distinguishable from each other as well. The colour scheme has been changed to a rainbow scheme, which shows up the differences more distinctly. Use of a log scale has also been used where necessary.

"Section 3.1: I found this section far too detailed. Maybe some of the content could be moved to the supplementary material?"

We agree. This section has been moved to the supplementary material.

"*Fig6: Please provide a legend for symbols and colours within the figure.*"

Legends for the symbols and colours used in all plots have been provided, instead of including this information in the caption. This particular figure has been dropped, and replaced with an expanded time series where day and night concentrations have been separated.

"*Fig7: Is there a reason why Oct12, Dec12, Jan13 are missing? I am not sure if the diurnal cycle for all months is required in the manuscript?*"

These are plots of the observed diurnal pattern in the concentrations. Unfortunately observations were not available for all months, and there were large gaps on the concentration data during these three months. The monthly plots have been moved to the supplementary material, and instead the average diurnal plot over all periods has been improved in the manuscript (Figure 8).

"*Table 2: Consider to move this table to supplementary material.*"

Agreed. This will be moved together with Section 3.1. to the supplementary material.

"*P34, L8-L9: NEE or NEP?*"

The net primary productivity was used as the uncertainty in the net ecosystem exchange. This has been corrected in the manuscript. Generally, for most parts of South Africa, the NEE flux itself is very small, but could be resulting from relatively large respiration and photosynthesis fluxes. If the uncertainty in the NEE estimate was based on the NEE value itself, it would be unrealistically small. This discussion has been made clearer in the revised manuscript.

This paragraph has been removed in the revised manuscript and these details now appear in section 2.9.2. "The uncertainty in the biogenic prior fluxes was set at the absolute value of the net primary productivity (NPP) as produced by CABLE. This is a very large error relative to the prior estimate itself, but there is a great deal of uncertainty in both the productivity and respiration fluxes contributing to thee NEE flux (Wang et al. 2011)."

"*Figs. 9+10: Please provide a legend for symbols and colours within the figure.*"

These figures have been changed altogether to ensure that the time series is not squashed, and to separate out the day and night concentrations and residuals. Better use has been made of legends. Please see response to Referee 1 for an example of the new time series plots. These are now figures 6 and 7.

"*Fig.12: Please label the sub-plots (i.e. north, east, south, west)*"

Thank you. This has been done. This is now Figure 11.

"*P42, L1-L3: Please rephrase this sentence.*"

This sentence has been revised.

Previous: "The innovation of the inversion can be observed through the differences between the prior and posterior flux estimates, which we refer to as the innovations, and through the change in the uncertainty estimates."

Revised: "We refer to the difference between the prior and posterior flux estimates as the innovations. The impact of the inversion on the flux estimates can be assessed through the size and direction of these innovations and through the reduction in the flux uncertainties."

"*Figs.13+14: Add labels and improve colour scheme.*"

These plots have been improved. As discussed earlier, the colour scheme was selected to ensure that positive, negative and zero fluxes were distinct. The jagged colour graduation was also selected so that large fluxes could be displayed together with small fluxes, while still showing subtle differences between the smaller fluxes. The colour scheme has been changed to the rainbow scheme to make differences between pixels more distinctive. These are now Figures 12 and 13.

"*Table 3: Is this level of detail really required in the manuscript?*"

The purpose of including this table was so that the impact of the inversion could be assessed from month to month. This table attempts to summarise the distribution of the fluxes over the domain, and show how that distribution is changed by the inversion. This section has been largely rewritten. The table has been converted into a time series of box plots to show the distribution of these pixel-level fluxes (Figure 14).

"*P49: I have difficulties understanding the whole section on this page.*"

This section has been largely rewritten. The main point that this section makes is that, under the current inversion framework, the posterior off-diagonal covariances are very small. This is because the uncertainties in the modelled concentrations that are attributed to the flux contributions ($\mathbf{HC_{s_0}H}^T$) are small relative to the uncertainties specified in the observation error covariance matrix ($\mathbf{C_c}$). If the diagonal elements in $\mathbf{C_c}$ could be made smaller, then the posterior covariances would be much larger in magnitude. When we wish to estimate the difference between the fossil fuel and NEE fluxes $s_{f,i} - s_{NEE,i}$, in other words, to distinguish between the fossil fuel and NEE fluxes in the same pixel, the variance of this estimate is determined by the sum of the variances of these two fluxes minus twice the covariance between them: $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} - 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ will be negative. If the posterior covariances were larger in magnitude (resulting in large negative correlations), then the inversion would be less able to distinguish between these two fluxes. At the moment the dominant term is the large uncertainty prescribed to the NEE estimates. Therefore, even though the negative correlations are small, the posterior variance of $s_{f,i} - s_{NEE,i}$ is still large, therefore the inversion does not distinguish well between NEE and fossil fuel fluxes under the current configuration.

But when we aggregate the fluxes from the same pixel, the variance of NEE plus the fossil fuel flux from the cell would be $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} + 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ will be negative. The inversion always results in reductions in the uncertainty of the individual fluxes (or at worst they remain the same). If posterior covariances between the flux uncertainties were larger, which would occur if the uncertainties prescribed to the prior fossil fuel and NEE fluxes were smaller or if the elements of the observation error covariance matrix were smaller, then there would be a much smaller

uncertainty in the aggregated flux estimates. The negative posterior covariances which are produced by the inversion mean that the uncertainty of the aggregated flux is always smaller than the sum of the uncertainty of the individual components. But by how much smaller depends on the prior uncertainties and the skill of the atmospheric transport model.

This discussion appears on page 45 of the revised manuscript, at the end of section 3.2.

"*Figs. 15+16: I cant see the value of those figures. I don't think they are even referenced in the text?*"

Only Figure 15 is retained (which is Figure 15 in the revised manuscript). It was an oversight that these were not originally reference, and this has been corrected. The discussion has been expanded as explained above. See also the response to Referee 1 on their comment marked B.

"*Section 3.2.3: Can you please check the units? You refer to fluxes, but units are in ktCO2.*"

The units are $\mathrm{kt\ CO_2\ week^{-1}}$ emitted from the domain of the inversion. This has been clarified.

"*Fig. 17: Please add a colour legend. Check units.*"

This has been clarified and legend added. This is Figure 16 in the revised manuscript.

"*Section 3.2.4: Units not consistent. I think, they are all fluxes.*"

This section has been merged with section 3.2.3 which now considers the aggregated fluxes. The units in the original manuscript were $\mathrm{kt\ CO_2\ month^{-1}}$. We now consider only the aggregated weekly fluxes.

"*Table 5: move to supplementary material?*"

This whole section has been moved to the supplementary material.

References

Ziehn, T., Nickless, A., Rayner, P. J., Law, R. M., Roff, G., and Fraser, P.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 1: Methodology and Australian test case, Atmos. Phys. Chem., 14, 9363–9378, doi: 10.5194/acp-14-9363-2014, 2014.

Estimates of CO_2 fluxes over the City of Cape Town, South Africa, through Bayesian inverse modelling

Alecia Nickless[1,2], Peter J. Rayner[3], Francois Engelbrecht[4], Ernst-Günther Brunke[5], Birgit Erni[1], and Robert J. Scholes[6]

[1]Department of Statistical Sciences, University of Cape Town, Cape Town, 7701, South Africa
[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, OX2 6GG, UK
[3]School of Earth Sciences, University of Melbourne, Melbourne, VIC 3010, Australia
[4]Climate Studies and Modelling and Environmental Health, CSIR, Pretoria, 0005, South Africa
[5]South African Weather Service c/o CSIR, P.O. Box 320, Stellenbosch, 7599, South Africa
[6]Global Change Institute, University of the Witwatersrand, Johannesburg, 2050, South Africa

Correspondence to: Alecia Nickless alecia.nickless@phc.ox.ac.uk

Abstract. ~~The results of a high resolution Bayesian inversion over the City of~~ We present a city-scale inversion over Cape Town, South Africa~~, are presented, which used observations of~~. Measurement sites for atmospheric CO_2 ~~from sites~~ concentrations were installed at Robben Island and Hangklip lighthouses~~collected over a sixteen month period from March 2012 until June 2013. A Lagrangian particle dispersion model driven by the regional climate model Conformal Cubic Atmospheric Model (CCAM) was used to provide the sensitivities of the observations to the surface sources and boundary concentrations. This regional climate model was dynamically coupled to the CABLE (Community Atmosphere Biosphere Land Exchange) model, which provided prior estimates of the biogenic fluxes~~, located downwind and upwind of the metropolis. Prior estimates of the fossil fuel ~~emissions~~ fluxes were obtained from ~~an inventory analysis specifically carried out for this inversion exercise, making use of vehicle count data, population census data, fuel usage at industrial point sources, and aviation and shipping vessel counts. The inversion solved for the actual concentration measurements at each site, which was made possible by the use of~~ a bespoke inventory analysis where emissions were spatially and temporally disaggregated and uncertainty estimates determined by means of error propagation techniques. Net ecosystem exchange (NEE) fluxes from biogenic processes were obtained from the land atmosphere exchange model CABLE (Community Atmosphere Biosphere Land Exchange). Uncertainty estimates were based on the estimates of net primary productivity. CABLE was dynamically coupled to the regional climate model CCAM (Conformal Cubic Atmospheric Model), which provided the climate inputs required to drive the Lagrangian particle dispersion model. The Bayesian inversion framework included a control vector where fossil fuel and NEE fluxes were solved for separately.

Due to the ~~Cape Point background site to provide information on the boundaries, and was necessary due to the effect of topography on the atmospheric transport, affecting particularly the sensitivity of the Robben Island site to~~ large prior uncertainty prescribed to the ~~surface fluxes. Night-time observations were included, but allocated much larger errors compared to the daytime observations.~~

The NEE fluxes, the current inversion framework was unable to adequately distinguish between the fossil fuel and NEE fluxes, but the inversion was able to substantially improve the agreement between the modelled and observed concentrations, and able to better represent the diurnal cycle in the concentrations compared with the prior modelled concentrations. The mean bias in the modelled concentrations was reduced from -2.9 , with interquartile range -9.1 to 3.7, for the prior modelled concentrations, to 0.5, with interquartile range -1.5 to 1.5, for the posterior modelled concentrations at Robben Island, and from a bias of 2.4 in the prior modelled concentrations at the Hangklip site, with interquartile range -2.3 to 6.5, to a bias of 0.04, with interquartile range -1.1 to 0.8. The standard deviations of the posterior residuals at both sites were reduced to values below that of the observed concentrations.

The inversion solved for working week and weekend fossil fuel emissions, and weekly biogenic fluxes, each split into day and night contributions, for each month; therefore six surface sources per week within each of the 10,201 surface pixels. The inversion was also allowed to solve for each of the four boundary concentrations (north, east, south and west), but these were provided with tight constraints provided by the background site. The inversion tended to reduce fossil fuel emissions over all months. During the warmer, drier months, the inversion increased the biogenic fluxes, but reduced the biogenic emissions during the cooler, wetter months. The uncertainty reduction in the total estimate for the domain over each month ranged between 8.6 to 40.0%for the biogenic fluxes and between 0.4 to 16.4% for was able to obtain improved estimates of the total fluxes within pixels and across the domain. The median of the uncertainty reductions of the total weekly flux estimates for the inversion domain of Cape Town was 28%, but reach as high as 50%. At the pixel level, uncertainty reductions of the total weekly flux reached up to 98%, but these large uncertainty reductions were for NEE-dominated pixels. Improved corrections to the fossil fuel fluxes . Model assessment by means of the χ^2 statistic indicated that the mean statistic was 1.48 over all months, indicating that either the prior values for the model errors or the uncertainty in the fluxes was not specified high enough for some months. A companion paper on sensitivity analyses will address different options for the specification of the correlations between errors in the modelled concentrations, how these prior errors are determined, how correlations are determined between the prior fluxes, and how the state vector is specified. Greater confidence is given to the inversion's ability to correct the total flux within each pixel, rather than the individual flux estimateswould be possible if the uncertainty around the prior NEE fluxes could be reduced. In order for this inversion framework to be operationalised for monitoring, reporting and verification (MRV) of emissions from Cape Town, the NEE component of the CO_2 budget needs to be better understood. Additional measurements of $\Delta^{14}C$ and $\delta^{13}C$ isotope measurements would be a beneficial component of an atmospheric monitoring programme aimed at MRV of CO_2 for any city which has significant biogenic influence, allowing improved separation of contributions from NEE and fossil fuel fluxes to the observed CO_2 concentration.

1 Introduction

Cities are under pressure to reduce their carbon dioxide emissions. In the last 10 years (2006 to 2015), the mean annual increase in carbon dioxide (CO_2) concentrations in the global atmosphere has been 2.11 ppm per year (Dlugokencky and Tans, 2016) (NOAA/ESRL 2016), a sharper rise in CO_2 emissions than the preceding decades (IPCC, 2014). ~~Of~~ Approximately 76% of current anthropogenic greenhouse gas emissions ~~,~~ are comprised of CO_2 ~~contributes 76%~~ contributions (IPCC 2014). While cities cover a mere 2% of the global land surface area, they are responsible for 70% of anthropogenic greenhouse gas emissions (UN–Habitat, 2011), and between 71 and 76% of CO_2 emissions from global final energy use ~~(Seto et al 2014)~~(Seto et al., 2014). Annual urban CO_2 emissions are more than double the net terrestrial or ocean carbon sinks (Le Quéré et al., 2013).

South Africa is the single largest emitter of CO_2 on the continent of Africa, and the 13th largest emitter in the world (Boden et al., 2011). South African cities are home to 63% of the present population (Statistics South Africa, 2011), and by 2030 this is predicted to be 71%. The ~~City~~ population of Cape Town (CT) ~~saw its population increase from 2,563,095 in 1996 to 3,740,026 in 2011, an overall increase of 46% (City of Cape Town, 2011). Although cities are by far the largest contributors of anthropogenic emissions, they are also~~ has been rising at 2.5% per annum over the past two decades, and currently is nearly 4 million (City of Cape Town, 2011). Cities are seen as having the greatest potential to provide solutions for emissions reduction and climate change mitigation (Seto et al., 2014; Wu et al., 2016). By ~~mitigating~~ reducing the CO_2 impact of cities, cities play a pivotal role in decreasing their own climate vulnerability. But there are also additional co-benefits which include improving air-quality, energy access, public health, city liveability, and developing the economy and job creation through advances in green technology (Seto et al., 2014).

Formal climate action plans are developed by governments and city managers ~~,~~ whereby the roadmap for implementing greener policies is provided, such as encouraging and developing public transport which makes use of low emission technologies, mass and rapid transport systems, and building retrofits (Sugar and Kennedy, 2013; Erickson and Tempest, 2014). Many cities are taking it on themselves to respond to the climate crisis, reacting to limited international and national policy progress ~~, which is viewed to be moving too slowly to address the required need for mitigation against climate change (Hutyra et al., 2014). To~~ (Hutyra et al., 2014). But to determine if the plans implemented are having the anticipated effect of lowering CO_2 emissions, monitoring is required. ~~Emissions need to be known at baseline, and monitored through time.~~ Monitoring, reporting and verification (MRV) is a concept which is fundamental to most market and policy-based mechanisms in climate economics (Bellassen and Stephan, 2015). In order for emission reduction strategies to be properly implemented and assessed, an MRV approach should be adopted so that emission reduction claims can be validated in a consistent and reliable manner. Currently, the primary source of this information for cities is by means of emissions inventories. This relies on the collection of activity data to provide an inventory of emissions from different sectors or specific point sources. These inventories are not perfect representations of CO_2 emissions~~,~~. They are heavily dependent on accurate reporting, emission factors, and on assumptions regarding temporal or spatial disaggregation of emissions (Andres et al., 2012)~~. As the resolution of the inventory analysis increases, so too do the~~ , where errors associated with these emission estimates increase with higher spatial

and temporal resolutions (Andres et al., 2014). As the importance of these inventories increases ~~with~~ due to the need to quantify emissions and assess emission targets, it has become necessary to verify the accuracy of these estimates (NRC, 2010). Adequate MRV implementation requires transparency, quality and comparability of information, with narrow uncertainty estimates (Wu et al., 2016). Currently, uncertainties associated with urban emissions far exceed emission reduction goals, and therefore verification remains challenging. The large amount of uncertainty is due to factors such as incomplete data, inconsistency in reporting between different institutions or facilities, fugitive emissions from point sources such as those caused by gas leaks, and methodology which is rarely checked against scientific standards and procedures (Hutyra et al., 2014). A way of verifying inventory data for a city, and reducing uncertainty of inventory estimates, is by means of the Bayesian atmospheric inversion technique.

~~Originally implemented to determine global, large scale sources and sinks of~~ This method aims to take advantage of continuous measurement of CO_2 ~~(e. g. Chevallier et al. (2010)),~~ concentrations from a network of atmospheric monitoring sites located in and around a city. By attempting to model the CO_2 concentrations at these sites, the inversion is able to provide corrections to the inventory of CO_2 emissions from the city, so that the mismatch between the modelled and observed concentrations is reduced.

Several regional or mesoscale ~~scale atmospheric inversions are becoming more common~~ atmospheric inversions have been published (Lauvaux et al., 2008, 2009, 2012; Schuh et al., 2013), and more recently ~~city scale inversion studies are being~~ city-scale inversion studies have been conducted in Europe and North America ~~(Strong et al., 2011; Duren and Miller, 2012; McKain et al.,~~ These top down approaches make use of an atmospheric transport model to relate observations of CO_2 concentrations in the atmosphere to the ~~emissions~~ CO_2 fluxes from the domain of interest (Lauvaux et al., 2012). This method applies corrections to the inventory data, which enters the inversion calculation by means of the prior estimates. This paper reports the results for ~~such an inverse study for Cape Town~~an atmospheric inversion for CT, South Africa.

~~Atmospheric top down approaches to determining~~ Making use of point measurements of CO_2 ~~emissions have the simultaneous advantage and disadvantage of capturing information from all sources and sinks, some of which may have been excluded from the inventory analysis of the domain. All emissions~~ concentrations means that the effects of all fluxes of CO_2 are observed as an aggregated total~~, therefore all emission sources are accounted for, but it~~ . It is challenging to separate out these aggregated CO_2 ~~emissions~~ fluxes into different components of the total CO_2 budget without additional measurements, such as $\Delta^{14}C$ (Turnbull et al., 2015) and $\delta^{13}C$ isotope measurements (Newman et al., 2016)~~or confidence about~~ , or without high confidence in the spatial and temporal patterns of ~~emission (Shiga et al., 2014). At the moment background~~ fluxes (Shiga et al., 2014). Even when additional measurements of CO_2 mole fractions are available, at the current point in time, background atmospheric conditions are not sufficiently characterised ~~in order~~ to use isotope tracers to ~~differentiate~~ discriminate between fossil fuel and biogenic ~~sources, as these measurements are far rarer than atmospheric measurements of mole fractions (Turnbull et al., 2015). Even then it would not be possible to assign anthropogenic emissions to each sector. Therefore it is necessary to conduct an atmospheric inversion study in conjunction with~~ fluxes (Turnbull et al., 2015). To conduct a Bayesian atmospheric inversion at the city-scale, a detailed CO_2 inventory analysis ~~if no such inventory exists~~is required, where all the main contributors to the anthropogenic CO_2 budget are considered. ~~This analysis underpins the assumptions of human behaviour driving the~~

anthropogenic emissions. It is important to verify these assumptions regarding human activity in order to assess if mitigation interventions are having the desired impact (Strong et al., 2011). Better Apart from their use in an atmospheric inversion, better understanding of the underlying processes at the urban scale and improved quantification will provide of CO_2 emissions provides information contributing towards the policy decisions made by urban practitioners, and helps to improve understanding of urban dynamicsand inform , and informs future scenarios (Hutyra et al., 2014). An example of this is the detailed street level inventory analysis undertaken in the Hestia project for U.S. cities Indianapolis, Los Angeles, Phoenix and Salt Lake City (Gurney et al., 2012; Davis et al., 2017). Preceding these inventories was the Vulcan inventory which covers the contiguous U.S. (Gurney et al., 2009). These detailed inventories have made possible atmospheric inversion exercisesor , as well as other top down methods for obtaining urban CO_2 flux estimates, for these cities (Strong et al., 2011; Brioude et al., 2013; Bréon et al., 2015; Lauvaux et al., 2016). Such a detailed inventory analysis is not available for any South African city, and therefore a detailed spatially and temporally disaggregated inventory analysis of direct CO_2 emissions was undertaken for CT specifically for the use of this atmospheric inversion exercise (Nickless et al., 2015a).

Atmospheric inversion models have their own inversions have various sources of uncertainty, which include atmospheric transport modelling errors (particularly at night when the planetary boundary layer is very shallow) shallow) (Geels et al., 2007); incorrect characterisation of prior errors flux estimates and their uncertainties (which includes errors from in the inventory analysis) (Bréon et al., 2015; Lauvaux et al., 2016); atmospheric measurement errors ; and representation errors (Gerbig et al., 2003); representation errors due to the comparison of a concentration measurement at a point with a modelled concentration representative of a surface grid box (Gerbig et al., 2003); and aggregation errors which occur as emissions fluxes from various sources are coerced into homogeneous grid cells across different data sources(Kaminski et al., 2001). In the case of cities, atmospheric transport modelling is further complicated by small-scale turbulence, highly heterogeneous surface characteristics, and urban heat island effects (Hutyra et al., 2014; Bréon et al., 2015).

Previous studies on estimating emission for cities have found that errors in atmospheric transport modelling are a significant contributor to the overall uncertainty of emission estimates (Lauvaux et al., 2013, 2016; Bréon et al., 2015), and therefore more work is required to refine these models so that they can perform more reliably during these periods of high uncertainty before they can be used to infer emission estimates at all times of the dayTherefore, careful consideration of the atmospheric transport model (or models) is required for an atmospheric inversion. The atmospheric transport modelling in this study is was provided by the Conformal Cubic Atmospheric Model (CCAM) (McGregor and Dix, 2008) at the resolution of $1\,\mathrm{km} \times 1\,\mathrm{km}$. This model CCAM, at a slightly lower resolutionwas previously coarser resolution, has already been used for a regional network design study over South Africa, making use of the same Bayesian inverse methodology a similar Bayesian inversion framework (Nickless et al., 2015b), and has been verified over southern Africa at relatively low resolutions through to ultra high resolution (South Africa and over the CT target region at a spatial resolution of up to $1\,\mathrm{km} \times 1\,\mathrm{km}$) (Engelbrecht et al., 2009, 2011)(Roux, 2009; Engelbrecht

To be able to verify emissions from underlying processes, higher resolution inverse modelling systems are needed to better understand and quantify emissions from different sectors. Lauvaux et al. (2016) considered High resolution inversions are required to quantify emissions down to the sector or point source level. Lauvaux et al. (2016) performed an ultra high resolution inversion where sector specific anthropogenic emissions were considered, but ignored biogenic emissionsfluxes. This was pos-

sible due to the selection of the dormant period ~~when biogenic emissions~~ for the inversion, when fluxes due the biosphere would have been at a minimum. When considering longer periods, or for cities in regions which may not have a dormant vegetation period, this assumption will not be valid, particularly for a medium-sized city, where natural processes can be a significant contributor to the carbon budget. Such would be the case for South African cities, such as CT and Johannesburg, where large national parks and other natural areas are located near or within city limits and within city vegetation growth is non-negligible. ~~Additionally, CT is surrounded by~~ CT is also surrounded by a large agricultural sector consisting of vineyards and fruit orchards. Ironically, there are features of cities which allow for better plant growth. For example, the urban heat island effect leads to a longer growing season for plants, and reduced wind within cities leads to less plant stress resulting in better plant growth (Buyantuyev and Wu 2012). In addition, nitrogen deposition within cities leads to increased nutrient availability, and particularly in arid regions, cities cause augmented water availability for plants (Hutyra et al., 2014). If allowed growing space, plants can make a significant contribution to the carbon budget of a city.

Therefore, for a city like CT, biogenic ~~emissions should not~~ fluxes cannot be ignored, and within atmospheric inversion studies are usually ~~accounted for~~ estimated by means of a land surface exchange model (Bréon et al., 2015; Staufer et al., 2016). Bréon et al. (2015) and Staufer et al. (2016) made use of the C-TESSEL land atmosphere scheme which is used in the ECMWF forecasting system. In this study we have made use of the CABLE (Community Atmosphere Biosphere Land Exchange) model to represent the biogenic CO_2 fluxes in the CO_2 budget (Kowalczyk et al., 2006). CABLE ~~was dynamically coupled to CCAM, so that the land-atmosphere exchange model was run at the same resolution as meteorology.~~ had the same spatial and temporal resolution as the meteorology. The average weekly fluxes for each pixel were calculated and used as the prior biogenic fluxes.

We present ~~the results of an atmospheric inversion which was used to model fluxes over the CTusing atmospheric concentration data that was collected over~~ a Bayesian inversion framework used to obtain estimates of CO_2 fluxes over CT, and present the results of the reference atmospheric inversion for a sixteen month period ~~.~~ from March 2012 until June 2013. The domain considered was a $100\,\mathrm{km} \times 100\,\mathrm{km}$ region with CT at the centre. The spatial resolution of the atmospheric transport model was set at $1\,\mathrm{km} \times 1\,\mathrm{km}$, and the ~~surfaces sources were set~~ spatial resolution of the surface fluxes was made to match this ~~spatial~~ resolution. Fluxes were solved for at a weekly time step, separately for day and night~~, with fossil~~ . Fossil fuel and biogenic fluxes ~~solved separately~~ were solved for separately, and fossil fuel fluxes separated ~~by~~ into week and weekend fluxes.

2 Methods

2.1 Bayesian Inverse Modelling Approach

The Bayesian synthesis inversion method, as described by Tarantola (2005) and Enting (2002), was used to solve for the ~~sources~~ fluxes in this study. This method has been ~~employed~~ described for global inversions (Bousquet et al., 1999; Kaminski et al., 1999; Rayner et al., 1999; Gurney et al., 2002; Peylin et al., 2002; Gurney et al., 2003; Law et al., 2003; Baker et al., 2006; Rayner et al., 2008; Ciais et al., 2010), as well as for many of the ~~city scale~~ recent city-scale inversions (Lauvaux et al., 2016; Bréon et al., 2015). ~~This methodology was employed over South Africa in a previous optimal network design study~~

~~(Nickless et al., 2015b).~~ The observed concentration (c) at a measurement station at a given time can be expressed as the sum of different contributions from the surface fluxes, from the ~~boundaries~~ domain boundaries, and from the initial concentration at the site. ~~A linear relationship can be used to describe the relationship between the modelled concentrations and the contribution from the sources (surface fluxes and boundary inflow)~~Concentrations at the measurement site can be modelled as:

$$\mathbf{c}_{mod} = \mathbf{Hs} \tag{1}$$

~~The vector of~~ where $\mathbf{c}_{mod}$ are the modelled concentrations ~~$\mathbf{c}_{mod}$ is a result of the contribution from the sources~~ and $\mathbf{s}$ ~~, described by the transport or sensitivity matrix $\mathbf{H}$~~ are various sources, where sources are any part of the domain which can provide a positive or negative contribution of CO_2. $\mathbf{H}$ is the Jacobian matrix representing the first derivative of the modelled concentration at the observational site and dated with respect to the coefficients of the source components (Enting, 2002). ~~If we assume a Gaussian error distribution for the surface fluxes and concentrations we obtain the following cost function for our least squares problem:~~

It provides the sensitivity of each observation to each of the unknown sources, where the sources can be either fluxes or concentrations of CO_2. Estimates of the unknown sources can be obtained by minimising the following cost-function with respect to s:

$$J(\mathbf{s}) = \frac{1}{2}\left((\mathbf{c}_{mod} - \mathbf{c})^T {\mathbf{C}_\mathbf{c}}^{-1}(\mathbf{c}_{mod} - \mathbf{c}) + (\mathbf{s} - \mathbf{s_0})^T {\mathbf{C}_{\mathbf{s_0}}}^{-1}(\mathbf{s} - \mathbf{s_0})\right) \tag{2}$$

where ~~$\mathbf{C_c}$ is the error covariance matrix of the observations~~$\mathbf{s}$ is the control vector of unknown surface fluxes and boundary concentrations we wish to solve for, $\mathbf{s_0}$ is the vector of prior flux ~~estimates, s is the vector of predicted sources~~ and boundary concentration estimates, $\mathbf{C_c}$ is the uncertainty covariance matrix of the observations, and $\mathbf{C_{s_0}}$ is the ~~prior error~~ uncertainty covariance matrix of the ~~sources. The Bayesian cost function minimises both the difference between modelled concentrations and measurements and the difference between prior source estimates and predicted sources~~fluxes and boundary concentrations (Tarantola, 2005).

The ~~posterior source estimates can then be solved for as:~~ solution to this minimisation problem is:

$$\mathbf{s} = \mathbf{s_0} + \mathbf{C_{s_0}}\mathbf{H}^T\left(\mathbf{H}\mathbf{C_{s_0}}\mathbf{H}^T + \mathbf{C_c}\right)^{-1}(\mathbf{c} - \mathbf{H}\mathbf{s_0}) \tag{3}$$

~~The~~ and the posterior covariance matrix can be ~~calculated~~ determined as follows (Tarantola, 2005):

$$\mathbf{C_s} \equiv \left(\mathbf{H}^T \mathbf{C_c}^{-1} \mathbf{H} + \mathbf{C_{s_0}}^{-1}\right)^{-1}$$

$$\equiv \mathbf{C_{s_0}} - \mathbf{C_{s_0}} \mathbf{H}^T \left(\mathbf{H} \mathbf{C_{s_0}} \mathbf{H}^T + \mathbf{C_c}\right)^{-1} \mathbf{H} \mathbf{C_{s_0}}$$

$$\mathbf{C_s} \equiv \left(\mathbf{H}^T \mathbf{C_c}^{-1} \mathbf{H} + \mathbf{C_{s_0}}^{-1}\right)^{-1} \quad (4)$$

$$\equiv \mathbf{C_{s_0}} - \mathbf{C_{s_0}} \mathbf{H}^T \left(\mathbf{H} \mathbf{C_{s_0}} \mathbf{H}^T + \mathbf{C_c}\right)^{-1} \mathbf{H} \mathbf{C_{s_0}}. \quad (5)$$

2.2 ~~State~~ Control vector - s

The ~~state vectorwhich describes the total contribution to the concentration observed during an hour at a site from a particular surface source grid during~~ control vector, s, can be broken up into different components. The total CO_2 flux from a single surface pixel for a given week is ~~:~~ made up of the following individual fluxes:

$$\mathbf{s}_{sf;\,i} = \mathbf{s}_{\textbf{ff week day}}\mathbf{s}_{ff\ week\ day;\,i} + \mathbf{s}_{\textbf{ff week night}}\mathbf{s}_{ff\ week\ night;\,i} + \mathbf{s}_{\textbf{ff weekend day}}\mathbf{s}_{ff\ weekend\ day;\,i} + \mathbf{s}_{\textbf{ff weekend night}}\mathbf{s}_{ff\ weekend\ night;\,i} +$$

~~The surface source is made up of the fossil fuel emissions for a weekemitted~~ where $\mathbf{s}_{sf;\,i}$ is the total weekly surface flux from the i^{th} pixel, $\mathbf{s}_{ff\ week\ day;\,i}$ is the total fossil fuel flux during the day during the working week, $\mathbf{s}_{ff\ week\ night;\,i}$ is the total night-time fossil fuel flux during the working week~~and weekend separately for day and night,~~ , $\mathbf{s}_{ff\ weekend\ day;\,i}$ is the total weekend daytime fossil fuel flux, $\mathbf{s}_{ff\ weekend\ night;\,i}$ is the total weekend night-time fossil fuel flux, and $\mathbf{s}_{NEE\ day;\,i}$ and ~~the biogenic emissions~~ $\mathbf{s}_{NEE\ night;\,i}$ are the total day and night-time biogenic fluxes for the full week ~~, also separated into day and night fluxes. Each of these sources is solved for separately in the inversion solution. Finally,this total surface source contribution is added to the contribution from the mean boundary concentration for the week, which is solved for in the inversion rather than imposed, to give the observed concentration at the receptor site.~~

~~As the reference inversion which we present in this paper, we ran the inversion solution for the full month and solved for each week separately, four weeks in total. Therefore for a given surface source we solve for 2~~from the i^{th} pixel. The inversion solves for each of these separate fluxes. There are 101 $\times$ ~~4~~ 101 = ~~8 biogenic weekly fluxes(four day and~~ 10,201 surface pixels. Over the 16 month period from March 2012 to June 2013, separate monthly inversions are carried out for all months with sufficient valid concentration observations; a total of 13 inversions. Each monthly inversion solves for four ~~night) and~~ weekly fluxes. Therefore a monthly inversion solves for 10,201 $\times$ 6 $\times$ 4 = 244,824 surface fluxes.

The mean day and night-time concentrations at each of the four domain boundaries for each week are the last components of the control vector. The inversion solves for 4 $\times$ 2 $\times$ 4 = ~~16 fossil fuel fluxes (four day working week, four nightworking~~

week, four day weekend, and four night weekend) 32 boundary concentrations (4 boundaries, day/night, 4 weeks). We solved for weekly concentrations at the boundaries as we expected these concentrations to show small changes on synoptic time scales, particularly inflow from the ocean boundaries. We avoided solving for too short a period so that the percentile filtering technique (see section 2.8) would never discard all measurements for a period. The maximum standard deviation in the hourly background CO_2 concentrations for a week was 0.8 ppm.

As a sensitivity analysis, presented in a companion paper, we examined two alternative methods of solving for the weekly fluxcompositions of the control vector. We considered solving for a mean weekly flux for that each month. In this case for a surface source pixel we solved for two biogenic mean weekly fluxes (day and night) and four fossil fuel mean weekly fluxes (day and night working week, day and night weekend). Secondly we ran the inversion separately We also considered a separate inversion for each week. In this case only the concentration measurements for one week were used and the individual week weekly fluxes (two biogenic and four fossil fuel) were solved for, and this was repeated for each of the four weeks in the month, resulting in 24 individual fluxes solved for every month. The benefit of these two alternative methods provide control vectors is that the resulting dimensions of the $\mathbf{C}_{\mathbf{s_0}}$ matrix is much smaller compared to with the reference case we present in this paper.

In the reference case, for each month, the prior fluxes consisted of $101 \times 101 \times 4 \times 24 = 979{,}296$ surface sources for each month plus the $4 \times 2 \times 4 = 32$ boundary contributions, a large number of sources for which to solve compared to previous inversions of this kind. When solving for only one week, or a mean weekly flux for a particular month, the number of surface sources reduced to $101 \times 101 \times 24 = 244{,}824$. Solving for individual weeks required 4×2 additional boundary concentrations to be added to the state vector, and when solving for the mean weekly flux for the month, we allowed the boundary concentrations to differ for each week, and therefore we still solved for the 32 boundary concentrations as in the reference case.

2.3 Concentration measurements - c

The existence of the Cape Point Global Atmospheric Watch (GAW) station made CT an ideal candidate for a city scale inversion exercise. The Cape Point station is located approximately 60 south of CT within a nature reserve, situated on the southern-most tip of the Cape Peninsula at a latitude of $34°21'12.0''$ south and longitude of $18°29'25.2''$ east. The inlet is located on top of the 30 measurement tower, which is located on a cliff 230 above sea level. The station observes background measurements of when observing maritime air derived directly from the south-western Atlantic Ocean. This is an extensive region stretching from 20° (sub-equatorial) to 80° (Antarctic region) (Brunke et al., 2004). Therefore background measurements at Cape Point are well representative of the background signal influencing the Cape Peninsula. The background signal at Cape Point, obtained from a percentile filtering technique, was used as the prior estimate of background concentrations. The uncertainty in the background concentration is based on the hourly measurements contributing towards the weekly mean. As the variability in the background in the southern hemisphere is very small, much smaller than for the northern hemisphere, this results in a tight constraint on the prior background concentrations.

Two CO_2 monitoring sites were established at Robben Island and Hangklip lighthouses. Due to the dominant wind directions in CT (Fawcett et al., 2007), either from the south or north west, the location of the Robben Island and Hangklip stations were well suited for observing contributions from the area of interest, particularly from CT. One site would usually observe

~~mainly background influence whereas the other would be measuring~~ The Hangklip site observed mainly background air, but occasionally viewed the biogenic-influenced continental air. Robben Island often observed air with enhancements from CT. The location of these sites in relation to the domain are shown in Figure 1. The average wind speed and direction across the domain, as modelled by CCAM, are shown in the supplementary material (supplementary material section 1.3).

Figure 1. Google Earth image of the domain, where Cape Town is located at the centre. The corner coordinates of the full domain are 33°29′42.00″ south 18°11′42.00″ east (top left), 33°29′42.00″ south 19°12′18.00″ east (top right), 34°30′18.00″ south 18°11′42.00″ east (bottom left), 34°30′18.00″ south 19°12′18.00″ east (bottom left). The locations of the measurement sites and the Cape Point GAW station background site are indicated together with images of these sites (Photo credits: Ernst Brunke and Alecia Nickless). CBD = central business district.

Each site was equipped with a Picarro Cavity Ring-down Spectroscopy (CRDS) (Picarro G2301) instrument. This instrument measures CO_2, methane (CH_4), and water vapour (H_2O) simultaneously, every five seconds, producing a precision of better than 0.05 parts-per-million volume (ppmv) for CO_2, 0.07 parts-per-billion volume (ppbv) for CH_4, and 100 ppmv for H_2O. This instrument maintains high linearity, precision, and accuracy over changing environmental conditions, requiring only minimal calibration, and is recognised as ~~the~~ one of the highest precision instruments for measurement of the top three greenhouse ~~gasses~~ gases (Crosson, 2008).

The inlet of the measurement tube at each site was located at the top of the lighthouse, and had a Gelman filter to prevent contamination of the instrument through aerosols or water droplets. The inlet tube led to a VICI rotary valve which directed the sampled air stream to the Picarro instrument. Approximately every four days the rotary valve switched to a calibration line which allowed the flow of calibration gas through the instrument for a period of half an hour.

The Robben Island lighthouse is an 18 m tall circular masonry tower, and the height of the focal plane of the light is 47 m above the high water level. The location of the lighthouse is $33°48'52.20''$ south and $18°22'29.25''$ east. The Hangklip lighthouse is a 22 m tall concrete tower, where the focal plane of the light is 34 m above the high water level. It is located at $34°23'11.40''$ south and $18°49'42.30''$ east. It is located on the tip of False Bay, opposite to Cape Point.

2.4 System Meteorology

CCAM is the variable-resolution global atmospheric model developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (McGregor, 1996; McGregor and Dix, 2001; McGregor, 2005a, b; McGregor and Dix, 2008). It employs a semi-implicit semi-Lagrangian method to solve the hydrostatic primitive equations. The Geophysical Fluid Dynamics Laboratory (GFDL) parameterisations for long-wave and short-wave radiation are used (Lacis and Hansen, 1974; Schwarzkopf and Fels, 1991), with interactive cloud distributions determined by the liquid and ice-water scheme of Rotstayn (1997). Total-variation-diminishing vertical advection is applied to solve for the advective process in the vertical. A stability-dependent boundary layer scheme based on Monin Obukhov similarity theory is employed (McGregor, 1993), together with the non-local treatment of the boundary layer scheme as described in Holtslag and Boville (1993). A canopy scheme is included, as described by Kowalczyk et al. (1994), having 6 layers for soil temperatures and soil moisture (solving Richard's equation) and 3 layers for snow. The cumulus convection scheme uses a mass-flux closure (McGregor, 2003), and includes downdrafts, entrainment and detrainment. Gravity wave drag is parameterised following ~~(Chouinard et al., 1986)~~Chouinard et al. (1986).

CCAM may be applied in stretched-grid mode to function as a regional climate model, thereby providing a flexible framework for downscaling reanalysis data or global circulation model simulations to high resolution over an area of interest. Stretched grids are obtained using the Schmidt (1977) transformation. ~~Here the model is applied to downscale reanalysis data of the National Centres for Environmental Prediction (NCEP; (Kalnay et al., 1996)) to ultra-high resolution over the CT study area.~~ A multiple-nudging approach was followed to downscale the 250 km resolution ~~reanalysis data~~ National Centres for Environmental Prediction (NCEP) reanalysis data (Kalnay et al., 1996) to a resolution of 60 km over southern Africa, 8 km over the south western Cape and subsequently to a 1 km resolution over ~~Cape Point~~the study area. The 8 km resolution domain stretched over an area of about $1300 \times 1300\,km^2$, whilst the 1 km resolution domain centred over False Bay stretched over an

area of about $160 \times 160\,\mathrm{km}^2$. Output was stored at a time resolution of 1 hour. CCAM was spectrally nudged with the synoptic-scale forcing reanalysis data at 6-hourly intervals for the period 1979-2013 using a scale-selective Gaussian filter (Thatcher and McGregor, 2009, 2010). This forcing was applied from 900 hPa higher up into the atmosphere. Sea-surface temperatures from the NCEP data set were used as lower boundary forcing.

2.5 ~~Lagrangian Particle Dispersion Model (LPDM) - H~~

~~In~~ To justify the use of CCAM to provide modelled winds and other climatological variables, we rely on previous studies which have used this model for atmospheric transport modelling in our target area (Whittlestone et al., 2009), and studies which have validated CCAM at various spatial resolutions (Engelbrecht et al., 2009; Roux, 2009; Engelbrecht et al., 2011, 2013, 2015). In particular, CCAM has been able to satisfactorily recreate present-day rainfall totals and the rainfall seasonal cycle, as well as circulation patterns over South Africa (Engelbrecht et al., 2009), and has been able to simulate with some success mid-tropospheric closed-lows and extreme rainfall events (Engelbrecht et al., 2015). CCAM has been validated over the Stellenbosch wine-producing area, which falls within the domain of this inversion, with respect to temperature, relative humidity and wind speed at six different stations within this region (Roux, 2009). Those stations located within the high-resolution focus area of the stretched-grid obtained root mean square errors of 0.64 ms^{-1} or lower and correlations close to 1 between the modelled and observed wind speeds. Validating the wind product from CCAM further in a rigorous manner is beyond the scope of this paper.

2.5 Jacobian matrix - H

In order to generate the ~~sensitivity~~ Jacobian matrix, $\mathbf{H}$, for the inversion procedure, which maps the surface fluxes and boundary inflows to the concentrations observed at the receptor sites, a Lagrangian particle dispersion model (LPDM) ~~is~~ was run in backward mode. An LPDM simulates the release of a large number of particles from arbitrary ~~emissions sources~~ receptors and records the location of the particles at fixed time steps (Uliasz, 1993, 1994). The model implemented in this study was developed by Marek Uliasz (1993), which will be referred to as LPDM. LPDM is driven by the hourly three-dimensional fields of mean winds (u, v, w), potential temperature and turbulent kinetic energy (TKE), which ~~are~~ were obtained from the CCAM model. When LPDM is run backward in time, in receptor-orientated mode, the particle counts can be used to generate $\mathbf{H}$ for a given receptor site, as described in Ziehn et al. (2014) and Nickless et al. (2015b) following Seibert and Frank (2004).

~~We use the LPDM originally proposed by Uliasz (1994), which we run in reverse mode for each measurement station in the observation network. The sensitivity matrix for the four~~ The Jacobian for a 4 week period during each month of the study ~~is~~ was generated by allowing the LPDM model ~~run over a~~ to run in backward mode over a full 2 month period~~, and particle counts are~~. Particle counts were extracted for the 4 weeks of interest. Particles were released every 20 seconds and each particle's position was recorded at 1 minute intervals. Particles that were near the surface were allocated to a surface grid ~~cell~~ box, corresponding to the surface pixels of the atmospheric transport model, and the total particle count within each of these ~~was obtained to determine the surface influence or sensitivity~~boxes was determined. These counts depended on the dimensions and position of ~~these~~ the surface grid boxes. The particle counts were used to calculate the source–receptor (s–r) relationship~~, or influence~~

functions, which form the sensitivity matrix **T**. Here, we . We followed Seibert and Frank (2004) to derive the elements of that convert the particle counts into the elements of the Jacobian matrix. As described in Ziehn et al. (2014), we modified the approach of Seibert and Frank (2004) to account for the particle counts which were produced by our LPDM as opposed to the mass concentrations which were outputted output by the atmospheric transport model in their study. The resulting s–r relationship between the measurement site and source i at time interval n, which provide the elements of the matrix $\mathbf{H}$, is:

$$\frac{\partial \mathbf{c_{sf}}}{\partial \mathbf{s}_{in}} = \frac{\Delta T \mathrm{g}}{\Delta P} \overline{\left(\frac{N_{in}}{N_{tot}}\right)} \frac{44}{12} \times 10^3, \tag{7}$$

where $\bar{\chi}$ $\mathbf{c_{sf}}$ is a volume mixing ratio (receptor) expressed in ppm and $\dot{q}_{in}$ $\mathbf{s}_{in}$ is a mass flux density (source), N_{in} the number of particles in the receptor surface grid from source grid pixel i released at time interval n, ΔT is the length of the time interval, ΔP is the pressure difference in the surface layer, g is the acceleration due to gravity, and N_{tot} the total number of particles released during a given time interval.

In this inversion setup weekly fluxes of carbon CO_2 were separated into day and night-time contributions, into fossil fuel and NEE contributions, and in the case of fossil fuels, into working week and weekend contributions. The total flux for a source region is the sum of all these contributions, and therefore, the mass flux density $\dot{q}_{in}$ in Eqn. has units of C for the day and similarly for the night. In the case of the Therefore, to obtain the NEE contributions the particle count N_{in} is was the sum over one week (ΔT=1 week (day/night)). For fossil fuel sourcesfluxes, the particle count is was separated into the contribution from the working week and from the weekend, separately for day and night.

The surface layer height was set to $50\,\mathrm{m}$ which corresponds to approximately $595\,\mathrm{Pa}$ (ΔP). If we assume well mixed conditions, then the s–r relationship should be independent of the thickness of the surface layer, as long as the layer is not too deep, as the particle count will be adjusted proportional to the volume of the grid box. Under stable conditions, this may not be the case (Seibert and Frank, 2004).

The spatial resolution of the surface flux grid boxes was set to be the same as that of the high-resolution subregion of the atmospheric transport model, resulting in a grid gridded domain consisting of 101×101 grid boxes (a resolution of approximately $0.01^\circ \times 0.01^\circ$ or $1\,\mathrm{km}\times1\,\mathrm{km}$). As the land surface model CABLE was dynamically coupled to CCAM, the biospheric fluxes were at the same resolution as the atmospheric transport model. This high spatial resolution was deliberately maintained throughout the inversion to avoid any unnecessary aggregation error (Kaminski et al., 2001; Nickless et al., 2015b), and allowed for small scale transport features to be maintained in **H**.

Since the surface sources are expressed as fluxes of carbon, the contribution The fluxes from the surface pixels are expressed in $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ and are transformed through **H** into contributions to the concentration at the measurement site is expressed in the amount of carbon seen at the measurement site from a particular source. In the case of the boundary sources (or contributions from outside of the domain) which are given as concentrations , their contributions to the concentration at the measurement site are expressed as a proportion of their concentration, dependent on their influence at the receptor sitein units of ppm. The inversion solves for the concentrations at the boundary of the domain. Ziehn et al. (2014) shows that by calculating the Jacobian which provides the sensitivities of observed concentrations to boundary concentrations , the boundary

contribution can then can be calculated as:

$$\frac{\partial \bar{\mathbf{s}}_{\mathbf{B}}}{\partial \mathbf{c}_b} = \frac{N_B}{N_{tot}} \tag{8}$$

where $\bar{\mathbf{s}}_{\mathbf{B}}$ is the concentration at the domain boundary, $\mathbf{c}_b$ is the volume mixing ratio, N_B is the number of particles from the specific domain boundary and N_{tot} the total number of particles viewed at the receptor site from any of the domain boundaries. The contribution to the observed concentration at the receptor site can be written as:

$$\mathbf{c}_b = \mathbf{H}_B \mathbf{c} \mathbf{s}_B \tag{9}$$

where $\mathbf{H}_B$ is the Jacobian , $\mathbf{c}_B$ the with respect to the domain boundary concentrations, $\mathbf{s}_B$ the domain boundary concentrations and $\mathbf{c}_b$ the contributions from the boundary to the observed concentration at the measurement site in units of ppm. The row elements of $\mathbf{H}_B$ sum to one.

For the network design, four boundaries(north, south, east and west) were used and we calculated the sensitivity of hourly observed concentrations to weekly boundary concentrations (separate Therefore the elements of $\mathbf{c}_b$ represent a weighted average of the concentrations at the domain boundaries, and provide a basis concentration to which the contributions from the surface fluxes are added. Each inversion solves for 4 weekly domain boundary concentrations for each cardinal direction, separated by day and nightweekly concentrations).

2.6 Inventory of anthropogenic emissions

An inventory analysis was conducted specifically for this atmospheric inversion exercise (Nickless et al., 2015a). The population of CT was 3,740,025, as reported in the 2011 census (Statistics South Africa, 2011). At that time the energy use of the city was determined to be 50% due transport, 18% due to residential, 16% due to commercial, 14% due to industrial and 1% due to government use. But of the carbon emissions due to energy usage, only 27% were attributed to the transport sector as a result of the carbon intensive usage of coal for electricity generation to provide almost all of the energy to the residential and commercial sectors in South Africa, which emit approximately 29% and 28%, respectively, of the total carbon emissions of CT (City of Cape Town, 2011).

The anthropogenic emissions were subdivided into those due to road transport, airport and harbour emissions, residential lighting and heating, and industrial point sources. Road transport emissions were derived from modelled values of vehicle kilometres for each section of the road network, derived modelled from observed vehicle count data. The vehicle kilometres were scaled for each hour of the day, and reported separately for working week days and weekend days. Therefore the vehicle emissions for day and night are distinctive for the week / weekend and day / night periods.

Airport emissions were derived from landing and takeoff cycles, as reported by Airports Company South Africa for each month. We used the IPCC reported average emission factors for domestic and international fleets (IPCC, 2000), and these were

used to convert the airport activity data into emissions of CO_2. Emissions were expected to be concentrated ~~during the hours of~~ between 6:00 and 22:00, and so the monthly emission was divided evenly between these hours. Harbour emissions were derived for port activity published by the South African Ports Authority for each month. Based on the gross tonnage of vessels which docked at the port during the month, emissions could be derived as described in DEFRA (2010). The monthly emissions were divided equally between all hours of the month, as it was assumed that harbour activities would be continuous.

Residential emissions for lighting and heating were derived from population count data obtained for each of the municipal wards in 2011 (Statistics South Africa, 2011). The population of CT was 3,740,025, as reported in the 2011 census (Statistics South Africa, 2011). The South African government reports on the fuel used for domestic heating and lighting (South African Department of Energy, 2009). This was divided between the total population, and then allocated to each ward depending on the population residing in that area. The fuel usage was scaled according to the proportion of fuel used for cooking, lighting and heating, where 75% of the annual heating fuel usage was assumed to take place during the winter months (March to August). It was assumed that 75% of the annual energy consumed was used for heating, 20% for cooking and 5% for lighting.

CT provided monthly fuel usage by the largest industrial emitters. The reported fuel usage for the top fuel users were converted directly into CO_2 emissions by multiplying these figures with the Defra greenhouse gas emission factors (DEFRA, 2013a). The fuel types that were considered included heavy fuel oil, coal, diesel, paraffin and fuel gas which were divided into liquid petroleum gas and refinery fuel gas. As no information was available about when the activity was occurring at these facilities, the emissions were divided equally between all hours of the month.

Based on this inventory analysis, the percentage contribution of industrial point sources to the total fossil fuel emission for CT was 12.0%, 34.6% from vehicle road transport, 51.0% from the residential sector, and 2.4% from the airport and harbour transport. Residential emissions are a large contributor to the fossil fuel emission budget as well as one of the largest contributors to the uncertainties in the fossil fuel flux. This is due to the dependency that many people living in CT have on raw fossil fuel burning for heating and lighting. Emissions from power stations are a small component of the total fossil fuel flux from CT as the bulk of the direct emissions from power stations occur elsewhere in the country.

The total fossil fuel emissions for the domain were comparable with those from the EDGAR (Emission Database for Global Atmospheric Research) (v4.2) database (Nickless et al., 2015a). EDGAR is a global product on a $0.1^\circ \times 0.1^\circ$ grid, which provides the total anthropogenic emissions of CO_2 as estimated from proxy data such as population counts and information on the road transport network (Janssens-Maenhout et al., 2012). The total emissions estimated from our bespoke inventory analysis for 2012 were 22% higher relative to the emissions from EDGAR in 2010, where the emissions in our inventory tended to be concentrated over specific sources, such as oil refinery plants, whereas the EDGAR emissions were smoothed over the city region.

2.7 Biogenic emissions

CCAM was dynamically coupled to the land surface model CABLE, which allows for feedbacks between land surface and climate processes, such as leaf area feedback on maximal canopy conductance and latent heat fluxes (Zhang et al., 2013). This type

of coupling has successfully been implemented in CSIRO's national earth system modelling scheme (Australian Community Climate and Earth System Simulator or ACCESS) and describes land-atmosphere exchanges of energy, carbon, and water using biogeochemical, vegetation-dynamic and disturbance processes (Law et al., 2012). Several studies have validated CABLE under different ecosystems and parameters using both global model simulations (e.g. ~~Zhang et al. (2009)~~Zhang et al. (2009); Wang et al. (201
5 and site level offline CABLE simulations ~~(Zhang et al., 2013)~~(Exbrayat et al., 2013; Zhang et al., 2013).

The model produces hourly estimates of net ecosystem exchange (NEE), which were aggregated into weekly (day and night) ~~estimates~~flux estimates in units of $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$, and used as the prior estimate of biogenic fluxes over the land surface. The spatial resolution of these prior NEE fluxes were kept at a ~~1~~ $0.01^\circ \times 0.01^\circ$ resolution.

In terms of natural vegetation, the target domain is dominated by the fynbos biome. This biome is biodiverse, with many
10 endemic species, and covers a relatively small area in South Africa, but a significant area within the domain of the inversion. The fynbos biome is poorly represented by dynamic vegetation models (Moncrieff et al., 2015). The land atmosphere exchange model CABLE was selected to couple with CCAM due to its use and development in regions of Australia which share similar characteristics to the savanna biome in South Africa, which has a coverage of over 50%. Its ability to simulate respiration and photosynthesis in the fynbos region is largely untested. In addition to the natural vegetation, a large agricultural sector is within
15 the proximity of CT, consisting predominantly of vineyards and fruit orchards. The CT region experiences a Mediterranean climate with winter rainfall. Consequently, summers are hot and dry and winters are mild and wet. Therefore significant NEE fluxes take place during both winter and summer periods. The NEE in this region is limited by the amount of water availability, whereas temperatures are usually sufficiently high enough not to limit plant production and respiration.

The CO_2 fluxes over the ocean were obtained from Gregor and Monteiro (2013). This study characterised the seasonal cycle
20 of air-sea fluxes of CO_2 in the southern Benguela upwelling system off the South African west coast. A time series of pCO_2, derived from total alkalinity and dissolved inorganic carbon and scatterometer-based wind, was obtained from six monthly cross-shelf cruises in the St. Helena Bay region during 2010. Daily CO_2 fluxes were derived from these pCO_2. These fluxes were applied as prior estimates to the ocean surface grids within the domain. Therefore, an assumption was made that ocean CO_2 fluxes are relatively homogeneous in space near the south western coast of South Africa, but the inversion was given the
25 ability to differentially adjust each of the ocean sources in the posterior estimates.

2.8 ~~Boundary~~ Domain boundary concentrations

~~For this atmospheric inversion study, the selection of CT for this exercisewas determine by its location and the presence of~~ The existence of the Cape Point Global Atmospheric Watch (GAW) station made CT an ideal candidate for a city-scale inversion exercise. The Cape Point station is located approximately 60 km south of CT within a nature reserve, situated on
30 the southern-most tip of the Cape Peninsula at a latitude of $34°21'12.0''$ south and longitude of $18°29'25.2''$ east. The inlet is located on top of the 30 m measurement tower, which is located on a cliff 230 m above sea level. The station observes background measurements of CO_2 when observing maritime air advected directly from the south-western Atlantic Ocean. This is an extensive region stretching from 20° (sub-equatorial) to 80° (Antarctic region) (Brunke et al., 2004). Therefore, maritime measurements at Cape Point from the Southern Ocean are well representative of the background CO_2 signal influencing the

Cape Peninsula. The background signal at Cape Point, obtained from a percentile filtering technique (Brunke et al., 2004), was used as the prior estimate of the concentrations at each of the four domain boundaries. The percentile filtering technique removes data influenced by the continent or anthropogenic emissions. Two 11-day moving percentiles, which are adjustable by tuneable factors, control the upper and lower threshold limits. This results in a subset of background measurements from Cape Point represented by a narrow concentration band contained within these limits. This filter, when applied to the Cape Point ~~GAW station.~~ CO_2 measurements, selects approximately 75% of the data. The percentile-filtering technique has been shown to compare well with the more robust method of using contemporaneous radon (^{222}Rn) measurements to differentiate between marine and continental air.

This site provides a long term record of background CO_2 concentrations for the area~~, and provided the tightly constrained prior estimates of the background concentrations for the inversion~~. These continuous measurements of the background CO_2 levels meant that we were not dependent on the atmospheric transport model to produce ~~estimate of background~~ estimates of CO_2 ~~levels~~concentrations at the domain boundary, which are prone to large errors (Lauvaux et al., 2016). ~~This also meant that we did not have to depend on a gradient approachto estimate the fluxes, which results in measurements being used only when the wind direction and wind speed are within narrow limits (Bréon et al., 2015; Lauvaux et al., 2016; Staufer et al., 2016). The percentile-based statistical filter applied to the Cape Point~~ Due to the prevailing wind directions across the domain the "gradient approach" for solving for CO_2 ~~measurements to obtain background levels of atmospheric selects approximately 75% of the measurements (Brunke et al., 2004).~~ was not appropriate. This gradient approach relies on the observed wind direction and wind speed to obtain a subset of the concentration measurements when the air flow is from one measurement site directly to another. The differential in the concentrations is modelled by the inversion (Lauvaux et al., 2013; Bréon et al., 2015; Staufer et al., 2016). Plots provided in the supplementary material (supplementary material section 1.3) show the average wind speed and direction for the domain for each month. In general, the wind direction was not favourable to the gradient approach, and with only two measurement sites, would have left little information to constrain the surface fluxes. When the wind is blowing from the south easterly direction, air from the Hangklip site curves northwards towards the interior and away from CT. When the Robben Island site is observing marine air on its way into the CT area from the Atlantic side, such as June 2013, the wind changes for the north westerly direction once it passes over CT to a more northerly direction, missing the Hangklip site.

The mean weekly background ~~concentration~~concentrations, separate for day and night, ~~is~~ were determined from the percentile filtered measurements at the site, and ~~was~~ were used as the prior ~~boundary concentration~~ domain boundary concentrations for each of the four cardinal directions. The inversion was then allowed to ~~adjust these concentrationsslightly~~make small adjustments to these concentrations. The prior variance assigned to the boundary concentrations was equal to the variance of the measured hourly concentrations for that period. As the variability in the background CO_2 in the southern hemisphere is small, much smaller than for the northern hemisphere, this resulted in a tight constraint on the prior background CO_2 concentrations. Large adjustments by the inversion to the far-field domain boundary concentrations were not expected. The daytime weekly background concentrations are shown in Figure 2. The ~~uncertainty in the~~ standard deviation in the hourly background CO_2 concentrations ranged between 0.32 and 0.90 ppm, with a mean of 0.62 ppm.

The boundaries of the domain were deliberately set to be far from the measurement sites so that contributions to the CO_2 concentration at a measurement site were dominated by the surface fluxes within the domain, rather than by the domain boundary concentrations.

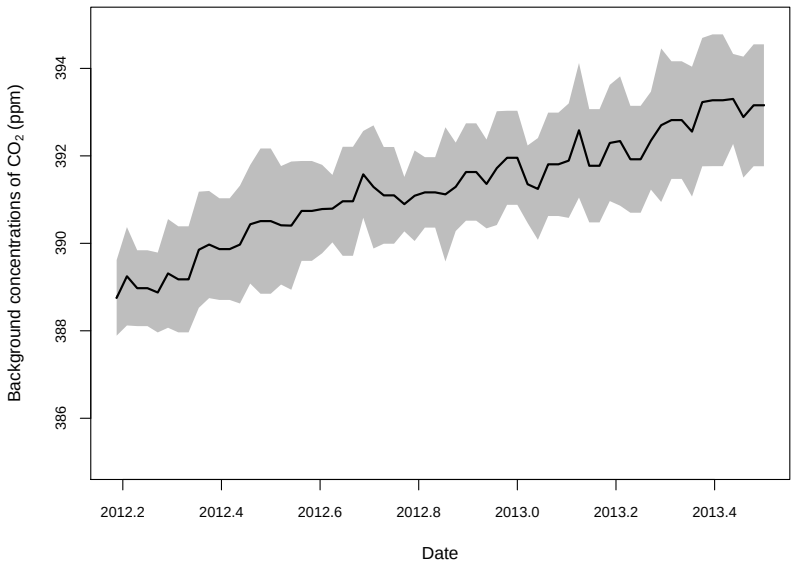


Figure 2. Weekly mean background concentrations of CO_2 (ppm) as measured at Cape Point GAW station, with 95% confidence interval represented by the grey shaded area. The mean concentrations are calculated from percentile filtered observations, extracting only those observations considered to be representative of background conditions.

2.9 ~~*A priori*~~ Prior covariance matrix - $\mathbf{C_{s_0}}$

The ~~variances and covariances in the *a priori* covariance matrixprovide the uncertainties in the prior estimates of the sources $\mathbf{s_0}$. The decision of whether to include non-zero covariances in~~ uncertainty covariance matrix, $\mathbf{C_{s_0}}$, of the prior fluxes and domain boundary concentrations $\mathbf{s_0}$ determines in part how much freedom the inversion has to ~~modify the source estimates~~ adjust these fluxes based on the observed concentrations $\mathbf{c}$. If the off-diagonal prior covariance elements are ~~significant~~ significantly different from zero, then the estimate for a ~~particular source~~ each flux will be more dependant on the prior estimates of the surrounding ~~sources~~ fluxes compared with an inversion where the ~~covariance terms are~~ covariances between the uncertainties in the prior fluxes were set to zero. ~~Covariance terms in the prior flux covariance matrix spread the influence of concentration measurements on the posterior flux estimates , subject to sensitivity matrix which determines which covariance terms contribute to the estimation of the flux.~~ On the other hand, if the prior variances are large, the inversion is able to make large adjustments to flux estimates to obtain better agreement between the observed and modelled concentrations. The next two subsections explain how the original estimates of the uncertainties in the fluxes and observation errors were determined. The uncertainties in the prior fluxes were scaled by an additional factor of 2 to ensure goodness-of-fit of the covariance structure (see supplementary material section 1.2).

2.9.1 Fossil fuel emissions

Error propagation techniques were used to estimate the uncertainties in the sector specific fossil fuel emissions. This ~~is~~ was described in Nickless et al. (2015a). ~~DEFRA (2013b) provides estimates of activity dataerror and emission factor error under each fuel type for industrial sources. These values were used to estimate the uncertainties in the industrial point source emissions using the following formula:~~

$$\delta \mathbf{s_0} = |\mathbf{s_0}| \times \sqrt{(\frac{\delta \mathbf{activity\ data}}{\mathbf{fuel\ usage}})^2 + (\frac{\delta \mathbf{emission\ factor}}{\mathbf{emission\ factor}})^2}$$

An industrial point source flux $s_{0;ff}$ was derived from the equation
~~In the case of the~~

$$s_{0;ff} = AE \tag{10}$$

where A is the activity data, usually fuel usage, and E is the process-specific emission factor. The uncertainty in the flux was estimated from

$$C_{s_{0;ff}} = |s_{0;ff}|^2 \times ((\frac{\delta A}{A})^2 + (\frac{\delta E}{E})^2) \tag{11}$$

where $C_{s_{0;ff}}$ is the uncertainty in the flux estimate expressed as a variance, δA is the uncertainty in the activity data and δE the

uncertainty in the emission factor, expressed as standard deviations. DEFRA (2013b) provides estimates of uncertainty in the activity data and emission factors under various industrial processes for each fuel type.

For vehicle emissions, which relied on count data, Poisson errors were assumed, and propagated together with the uncertainty in the conversion factors for the different vehicle types. For airport and harbour emissions, vessel counts were assumed to be correct, and therefore the uncertainty in the emissions contained within the emission factors for the different vessel types and activities. For aircraft, these errors are assumed to be 34% for the international fleet and 28% for the domestic fleet (IPCC, 2000). The error estimate for berth and manoeuvring activities of shipping vessels is reported to be between 20% and 30%, and therefore a conservative estimate of 30% was used (DEFRA, 2010). For domestic heating and lighting, the estimates relied on population census, which had a reported omission rate of 15%. There was no information available on the variability in fuel usage between households, and therefore the uncertainty in the domestic emissions was set at 30% as a relatively arbitrary, but conservative level. Domestic emissions due to fossil fuel burning was a large contributor to the overall fossil fuel flux of the domain. As the percentage uncertainty assigned to these fluxes was large, uncertainties in the domestic emissions was a significant contributor to the overall uncertainty in the fossil fuel fluxes.

~~The~~ After accounting for the scaling of the uncertainty estimates to improve goodness-of-fit of the covariance structure, the resulting uncertainty estimates (expressed as standard deviations) ranged between 6.7% to 71.7% of the prior fossil fuel emission estimate, with a median percentage of 34.9% to 38.4% depending on the month. These values are in general ~~slightly~~ more conservative compared ~~to~~ with uncertainties that were determined by Bréon et al. (2015) for the AirParif inventory, which were set at 20% throughout. The spatial distribution of the fossil fuel fluxes during the month of March 2012 are mapped in Figure 3. The daytime fossil fuel emissions have a mean of 0.006 $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$ and go up to 3.4 $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$. The mean went down to 0.004 $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$ during the summer months, when domestic heating and lighting fuel usage is lower. The largest fossil fuel emission estimated was located towards the north of the city, and corresponded to a crude oil refinery. Most point estimates were located on the outskirts of the city, with a few located within the central peninsula area. The road network is apparent in the figure of the prior fossil fuel fluxes displaying the corresponding transport emissions, and clearly illustrates the large contribution that road transport ~~contributes~~ makes to the overall $\mathrm{CO_2}$ budget of CT.

~~Prior estimates for day (top left) and night (top right) of the mean fossil fuel emissions () and the corresponding uncertainties, expressed as standard deviations () (bottom row), for the month of March 2012. These estimates were derived from an inventory analysis for CT based on vehicle, aviation and shipping vessel count data, population census data, and fuel usage at industrial point sources. These prior estimates are provided at a resolution of 1×1 and the extent of the grid is between 34.5° and 33.5° south and between 18.2° and 19.2° east.~~

~~Since we are solving~~ Since we solved for weekly, rather than daily fluxes, we ~~are using~~ used a strong assumption that fossil fuel fluxes within the same week ~~are~~ were 100% correlated. To allow the inversion to react to local conditions within a given week, no correlation ~~is~~ was assumed between weekly fluxes. Since fossil fuel emissions ~~are~~ were expected to be localised in space, we also assumed no spatial correlation ~~in the increments of the fluxes due to fossil fuel emissions. Relying on a correlation length approach to describe spatial correlation would be inadequate, as it could lead to correlation between locations with and without fossil fuel sources~~between fossil fuel fluxes.

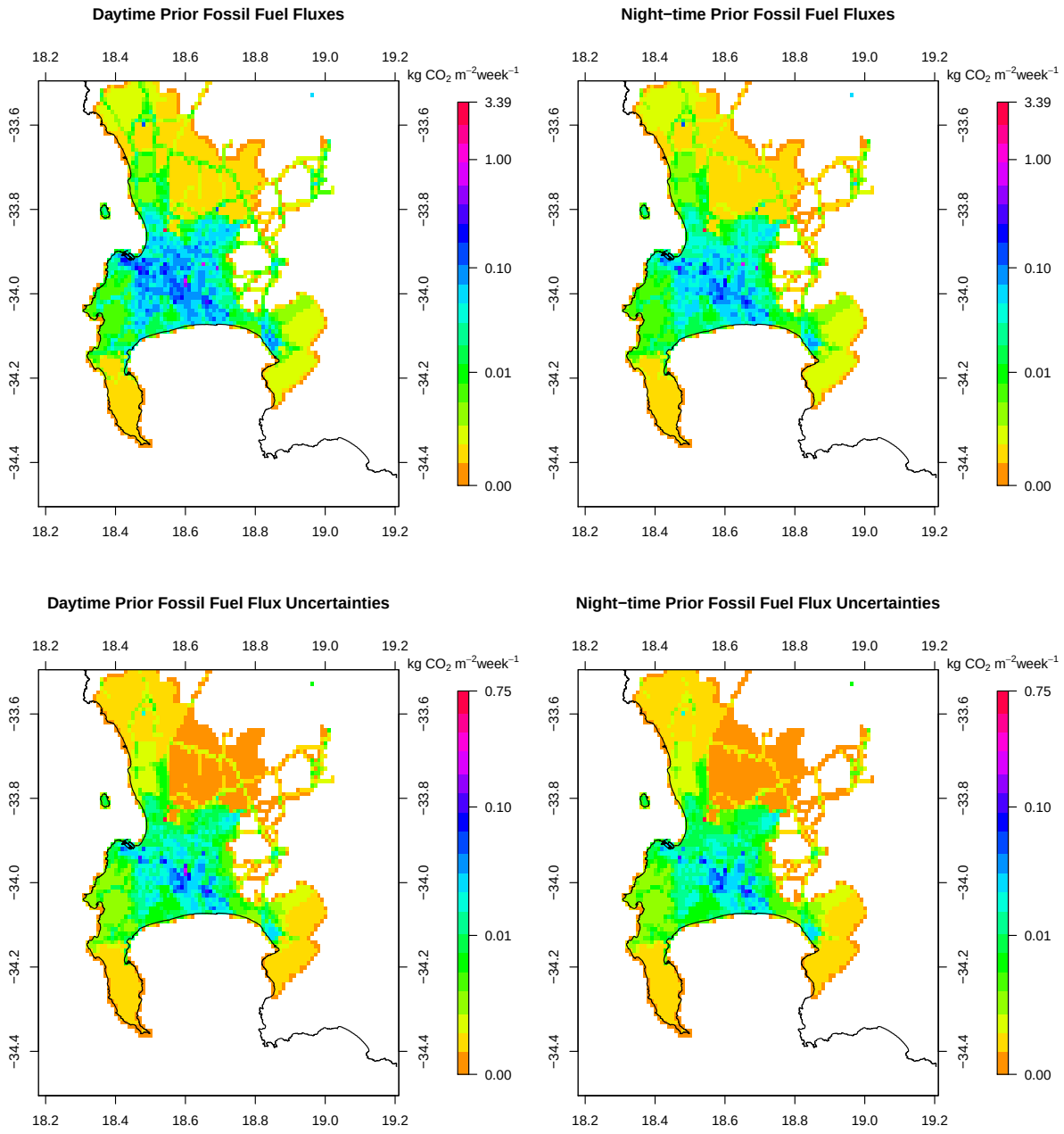


Figure 3. Prior estimates for day and night-time fossil fuel fluxes (kg CO_2 m^{-2} $week^{-1}$) and the corresponding uncertainties, expressed as standard deviations (kg CO_2 m^{-2} $week^{-1}$), for the month of March 2012. These estimates were derived from an inventory analysis for CT based on vehicle, aviation and shipping vessel count data, population census data, and fuel usage at industrial point sources. White indicates regions where the fossil fuel flux and its uncertainty are set to zero. These prior estimates are provided at a resolution of 1 km$\times$1 km and the extent of the grid is between 34.5° and 33.5° south and between 18.2° and 19.2° east.

2.9.2 Biogenic fluxes

The uncertainty in the ~~estimate of biogenic prior estimates~~ biogenic prior fluxes was set at the absolute value of the net primary productivity (NPP) as produced by CABLE. This is a ~~very~~ large error relative to the prior estimate itself, but there is a great deal of uncertainty in both the productivity and respiration ~~budgets for NEE. This approach is similar to that~~ fluxes contributing to the NEE flux (Wang et al., 2011). The estimates of NEE are strongly dependent on the assumptions behind the model forms selected for different processes in the CABLE model. For example, the model forms used for the soil temperature-respiration function and the soil moisture-respiration function have large impacts on the NEE estimates, with resulting NEE estimates differing by over 100% compared with measurements from flux towers (Exbrayat et al., 2013). The approach of assigning either the productivity or respiration component of NEE as the uncertainty has been used by Chevallier et al. (2010). ~~Rather than~~ We avoided assigning a fixed proportional ~~error, we preferred to base the uncertainty on the underlying processes driving NEE , as~~uncertainty to the NEE estimates as, particularly in semi-arid regions, such as those conditions found throughout South Africa, ~~very~~ small NEE fluxes can occur as a result of both large productivity and respiration fluxes. In the CT situation, this would lead to unrealistically low estimates of the uncertainty in NEE fluxes. This is different to the approach used by Bréon et al. (2015), where an uncertainty level of 70% was assigned to biogenic ~~emissions. In the CT situation, this would lead to unrealistically low estimates of the uncertainty in biogenic emissions~~fluxes, but in their case absolute NEE estimates were usually large in summer and expected to be small in winter.

~~We consider two parameterisations of the covariances between biogenic fluxes at different locations in space. For the reference inversion~~ To estimate covariances between the uncertainties in the NEE fluxes, we assumed an isotropic Balgovind correlation model as used in Wu et al. (2013). ~~As sensitivity analyses in a companion paper we considered no correlation in space as an alternative parameterisation. For the Balgovind correlation model, the~~ This helps to ensure positive-definiteness of the resulting covariance matrix. The off-diagonal covariance elements ~~between two spatial biogenic sources s_i and s_j are~~ for $s_{NEE;i}$ and $s_{NEE;j}$ were calculated as:

$$\mathbf{C_{s_0}(s_i, s_j)} C_{s_{0;NEE}}(s_{NEE;i}, s_{NEE;j}) = \sigma_b^2 \sqrt{C_{s_{0;NEE}}(s_{NEE;i})} \sqrt{C_{s_{0;NEE}}(s_{NEE;j})} (1 + \frac{h}{L}) exp(-\frac{h}{L}) \quad (12)$$

~~The~~ where $s_{NEE;i}$ and $s_{NEE;j}$ are NEE fluxes in pixels i and j, $C_{s_{0;NEE}}(s_{NEE;i})$ and $C_{s_{0;NEE}}(s_{NEE;j})$ the corresponding variances in the NEE flux uncertainties in pixels i and j, the characteristic correlation length L ~~is~~ was assumed to be 1 km, and h is the spatial distance between ~~s_i and s_j~~pixels i and j. As for the fossil fuel ~~sources~~fluxes, no correlation ~~is~~ was assumed between weekly biogenic fluxes, since the inversion setup is already assuming that biogenic fluxes within the same week ~~are~~ were 100% correlated ~~.~~ (i.e. constant over the week).

Figure 4 shows the spatial distribution of the ~~biogenic emissions~~ NEE fluxes and their uncertainties for the month of March 2012. Day-time ~~biogenic emissions range~~ NEE fluxes ranged between -0.19 and 0.04 kg CO_2 m^{-2} $week^{-1}$, concentrated over areas such as the Cape Point Nature Reserve and Kogelberg Nature Reserve, located near the Hangklip lighthouse. At night the ~~emissions are~~ fluxes were between 0.0 and 0.06 kg CO_2 m^{-2} $week^{-1}$. The uncertainties in the ~~biogenic daytime emissions have a median of 192% of the estimated emission and a mean of 881%, which translates to standard deviations ranging~~ NEE

daytime fluxes ranged between 0.00001 ~~(~~over the ocean) and 0.30 $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$~~. Depending on the month, the median uncertainty can range between 170 to 200% of the estimated biogenic emission, and estimates of biogenic emissions range~~ , whereas at night the uncertainties ranged between between ~~-0.22~~ 0.000001 and ~~0.004~~0.006 $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$. Uncertainties were smaller at night because night-time biogenic activity was mainly driven by respiration, and consequently the flux estimates were smaller as well as their uncertainties. Over the full measurement period, the estimates of NEE fluxes ranged between -0.22 and 0.004 during the summer to -0.11 to 0.007 $\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$ ~~at~~ during mid winter. ~~At night, when only the biogenic respiration process is in operation, the median uncertainty drops to 23% of the estimated emission and the mean down to 75%, which translates to standard deviations ranging between 0.000001 and 0.006 .~~

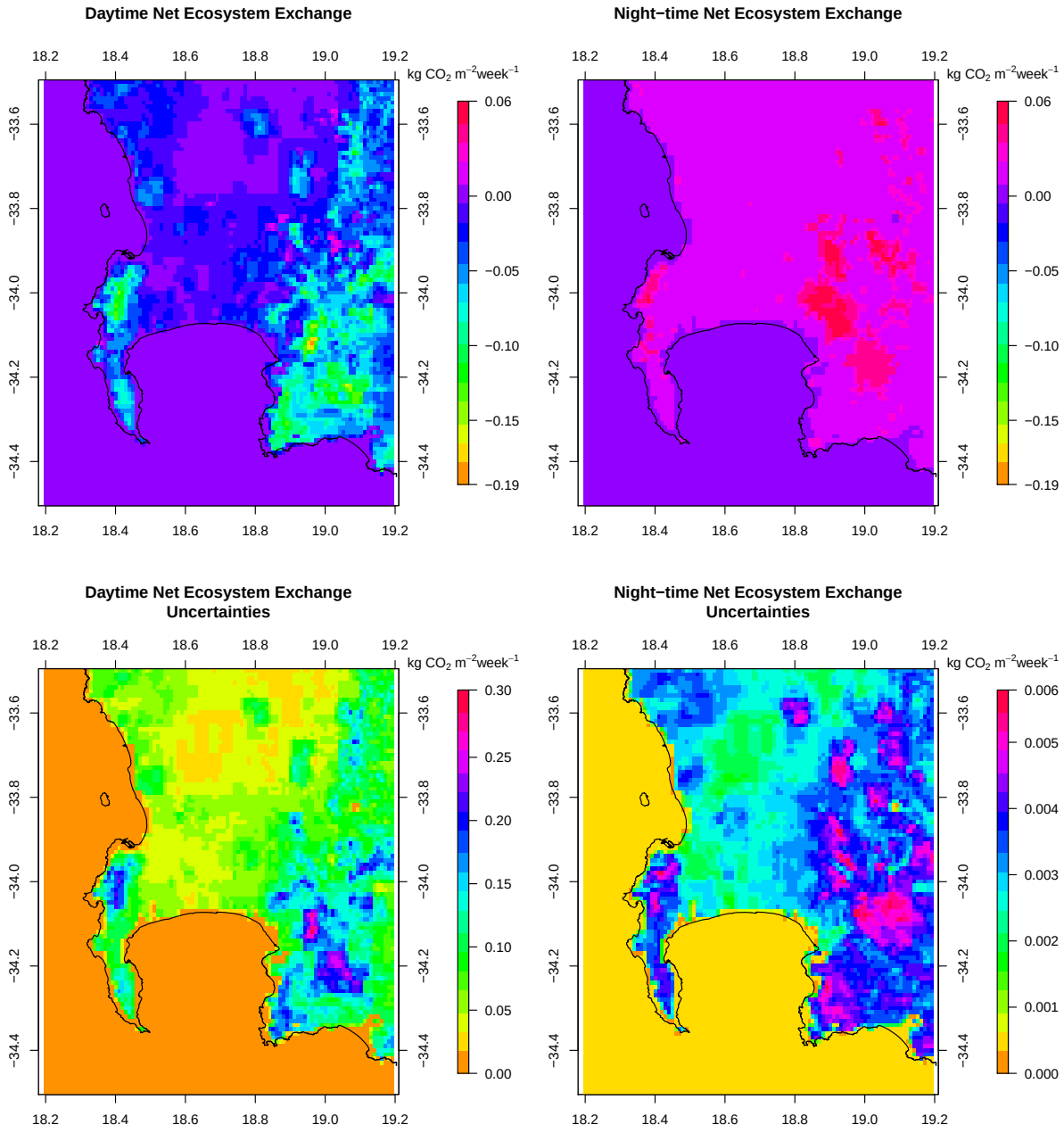


Figure 4. Prior estimates for day ~~(top left)~~ and ~~night~~ night-time NEE fluxes (~~top right) of biogenic emissions (~~$\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$) and the corresponding uncertainties, expressed as standard deviations ($\mathrm{kg\ CO_2\ m^{-2}\ week^{-1}}$)~~(bottom row)~~, during the month of March 2012. The prior estimates were obtained from the CABLE land-atmosphere exchange model ~~, which was coupled to the regional climate model CCAM and run~~ at a spatial resolution of $1\,\mathrm{km} \times 1\,\mathrm{km}$. The extent of the grid is between 34.5° and 33.5° south and between 18.2° and 19.2° east.

2.10 ~~Observation errors~~ Uncertainty covariance matrix of the observations - $\mathbf{C_c}$

The ~~observation errors~~ uncertainties in the observations represented in $\mathbf{C_c}$ contain both the measurement error (which are known to be in the order of 0.3 ppm) (Bréon et al., 2015; Wu et al., 2016) and the error ~~associate~~ associated with modelling the concentrations. The modelling errors result from several sources, including errors within the atmospheric transport model and aggregation errors which are due to smoothing emission estimates from localised sources within the spatial grids (Kaminski et al., 2001).

~~Several sources of error were considered for the observation error covariance matrix.~~ Similar to the approach adopted in the optimal network design for South Africa (Nickless et al., 2015b), ~~to account for transport modelling errors,~~ an error of 2 ppm during the day and 4 ppm at night was assigned to each observation, so that night-time observations carried less ~~influence~~ weight in the inversion. ~~We did not account any further for aggregation or representation errorsas we did in the network design, as we were running the inversion at the same spatial scale as the transport model. We acknowledge that there are still large sources of aggregation error due to smoothing of point source emissions across a relatively large grid, although we have kept the grid as small as could be managed in the inversion. The aggregation~~ These values were assigned as baseline (i.e. minimum) errors, and accounted for measurement errors, atmospheric transport modelling errors, aggregation errors and representation errors~~are therefore contained within the assigned 2 or 4 . These errors may seem smaller relative to city scale inversion~~ .

These errors are smaller than those for city-scale inversions conducted in the Northern Hemisphere. We justify the use of these values in our application since we are dealing with a much smaller city compared ~~to~~ with the megacity applications, such as Paris and Indianapolis. ~~The sources in this application would each result in smaller deviations from background compared to larger cities.~~ Measurements of background CO_2 have ~~also~~ shown that CO_2 concentrations in the Southern Hemisphere have smaller standard deviations. For example, for the years 2012 to 2013 the standard deviation between the monthly CO_2 means for Mauna Loa GAW station in the Northern Hemisphere was 2.3 ppm (Tans and Keeling , 2016), whereas for the same time period at Cape Point the standard deviation between the monthly means was 1.6 ppm.

~~As our reference approach, where we wanted to be as inclusive as possible of all sources of observation error , we~~ We accounted for additional sources of error in the atmospheric transport model. We took into consideration that ~~transport~~ errors in the modelled CO_2 concentrations due to the transport model would be larger when the wind speed was lower (Bréon et al., 2015), and this would be compounded at night when the planetary boundary layer height was lower and less stable (Feng et al., 2016). Additional error ranging between 0 and 1 ppm was added to the daytime uncertainty of 2 ppm, linearly scaled depending on the wind speed, with 0 ppm added when wind speeds were high (20 $\mathrm{m\,s^{-1}}$) and 1 ppm added when the wind speed was close to zero. At night the additional uncertainty ranged between 0 and 4 ppm.

We also considered the standard deviation of the measured CO_2 concentrations during each hour. It would be expected that if there was a large amount of variability between the instantaneous measurements at the site, ~~it is likely~~ that the atmospheric transport model would ~~have a greater chance of making~~ be more likely to make errors during this period. The ~~variability~~ variance of the observed CO_2 concentrations ~~, which~~ that contributed towards the mean estimate of the CO_2 concentration for that hour, was added to the overall ~~error. Therefore in the reference approach each hour has~~ uncertainty. Therefore each

hour had a customised observation error dependant on the prevailing conditions at the measurement site. Therefore the total observation error for hour k, as a variance, is given as:

$$C_c(ik,ik)^2 = E_{trans}{}^2 C_{c;base}{}^2 + E_{wind}{}^2 C_{c;wind}{}^2 + E_{obs}{}^2 C_{c;obs}{}^2 \quad (13)$$

where ~~E_{trans} is the transport modelling~~ $C_{c;base}$ is the baseline observation error of 2 ppm during the day and 4 ppm during the night, ~~E_{wind}~~ $C_{c;wind}$ is the additional error due to the wind speed conditions which ranged between 0 and 1, and ~~E_{obs} is variance~~ $C_{c;obs}$ is the standard deviation of the observed concentrations within that hour. A time series of the customised observation errors is provided in Figure 5. The final observation errors could reach up to 10 or 15 ppm at night, reducing the weight of these measurements in the estimation of the prior fluxes.

Temporal correlation between the observation errors was accounted for in an analogous manner to which covariance terms were estimated for the NEE flux uncertainties. The characteristic correlation length L was assumed to be 1 hour, and h was the temporal distance between observations.

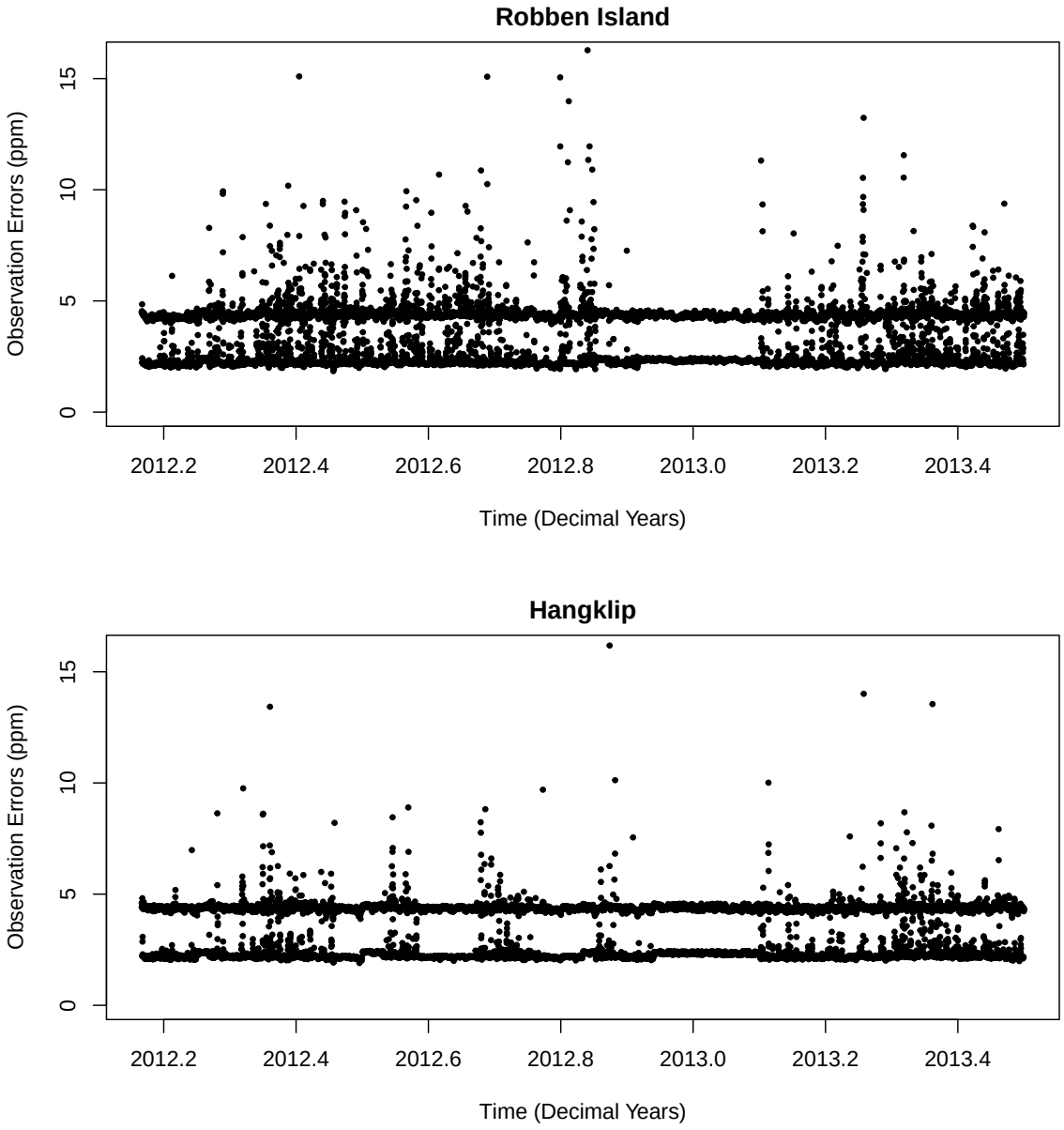


Figure 5. Time series of the customised observation errors (ppm) assigned to the CO_2 concentration measurement for each hour at the Robben Island and Hangklip measurement sites. The errors ~~are made up~~ consist of a ~~modelling~~ baseline error ~~component~~ (set as 2 ppm during the day and 4 ppm at night), ~~a~~ and additional atmospheric model errors based on prevailing wind speed ~~component,~~ and ~~measurement variability component~~the variation in the instantaneous CO_2 observations within an hour. The two distinct sets of points for each site arises due to the night-time ~~modelling error component~~ observation errors set to be ~~always~~ larger than ~~the~~ daytime ~~component~~observation errors.

2.11 Model Assessment

In order to assess the appropriateness of the ~~prior flux covariance parameters~~uncertainty covariance matrices $\mathbf{C_c}$ and $\mathbf{C_{s_0}}$, the χ^2 statistic, as described in Tarantola (2005) can be employed to determine the minimum value of the statistic:

$$\chi_1^2 = \frac{1}{\nu}(\mathbf{Hs_0} - \mathbf{c})^T(\mathbf{HC_{s_0}H}^T + \mathbf{C_c})^{-1}(\mathbf{Hs_0} - \mathbf{c}) \quad (14)$$

where ν is the dimension of the data space, in this case the length of observations in the inversion.

The squared residuals from the inversion (squared differences between observed and modelled concentrations) should follow the χ^2 distribution with degrees of freedom equal to the number of observations (Michalak et al., 2005; Tarantola, 2005). Dividing this statistic by the degrees of freedom should yield a χ_1^2 distribution. Values lower than one indicate that the ~~variance~~ uncertainty is too large, and values greater than one indicate that the ~~variance~~ uncertainty prescribed is lower than it should be~~and therefore the posterior estimates will be over-constrained by the prior fluxes.~~. The error in the assignment of the uncertainty could be in either $\mathbf{C_c}$ or $\mathbf{C_{s_0}}$ (or both).

Sensitivity analyses carried out on the specification of the covariance matrices have indicated that these errors are most likely contained in $\mathbf{C_{s_0}}$. These analyses are presented in a companion paper. In order to ensure the suitability of ~~the prior flux covariance matrix~~$\mathbf{C_{s_0}}$, the prior variances were ~~scaled~~ multiplied by a factor of two. This ensured that the χ_1^2 statistic was close to a value of one for almost all months of the inversion. A single scaling factor was used to adjust all the prior flux variances. An alternative to a single value scaling factor will be considered in a subsequent paper.

Using the χ^2 statistic to scale or estimate covariance parameters has been implemented by Lauvaux et al. (2016) and Michalak et al. (2005). Lauvaux et al. (2016) used the χ^2 statistic to scale the elements of the observation error covariance matrix. ~~As the estimation of the observation error covariance matrix is subject to a sensitivity analysis in the companion paper, the scaling was applied to the variance elements of the prior flux covariance matrix. Alternatively,~~ An alternative to manually scaling the elements of either $\mathbf{C_{s_0}}$ or $\mathbf{C_c}$, is to use a hierarchical Bayes approach ~~can be used~~ to estimate hyper-parameters for the covariance matrix, which are estimated based on the observed concentrations (Ganesan et al., 2014).

3 Results

The following section provides the results from the concentration measurements obtained at the two sites over the sixteen month period, followed by the results In this paper we concentrate on the results of the reference inversion. The extent to which the atmospheric transport model was able to represent the wind conditions at the sites is , as described in the supplementary material (Supplementary material Section 1.1).

3.1 Measurements

A single calibration standard was kept at each site, and run periodically in order to assess whether any drift was occurring in the measurements over time, to determine if the calibration coefficients required any adjustment. At Robben Island the measurements of the calibration standard were (mean ± (standard deviation)) 386.89 (± 0.014) in November, 385.67 (± 0.012) in February 2013, and 385.73 (± 0.012) in June 2013. This indicates that the instrument was making stable measurements during the sixteen month campaign. The slightly higher reading in November 2012 occurred when the weather was wet, and there was more moisture contamination from the previous ambient measurements prior to calibration. At Hangklip the instrument measured the calibration standard at 378.26 (± 0.009) in November 2012 and at 378.16 (± 0.022) in June 2013, indicating no evidence of drift over the measurement period.

In addition to each site's calibration standard, a travelling reference standard was also measured at close proximity in time at all three sites, including Cape Point. This standard consisted of clean air collected at the Cape Point site. The results of two of these calibration measurements are presented in Table **??**. By comparing calibration measurements with Cape Point inter-site differences are ensured to be negligible, and also links the Robben Island and Hangklip sites into the greater GAW network. The instruments at Cape Point are routinely calibrated with standards shared around the GAW network, which maintains high levels of quality control. Differences were found to be small between sites, and between calibration periods using the same standard in June 2012 and June 2013.

Mean and standard deviation of measurements () during calibration phase using a travelling standard during June 2012 and June 2013. The Cape Point site is considered the gold standard as the instruments here are maintained to be within GAW limits. FA01830Date Cape Point Robben Island previous section. We present sensitivity analyses elsewhere. Additional information on the distribution and time series of the observed concentrations at Robben Island, Hangklip Maximum inter-site differenceJune 2012 452.77 (0.03) 452.88 (0.02) 452.89 (0.03)0.12June 2013 452.28 (0.03) 452.86 (0.02) 452.60 (0.03)0.58June 2012 - June 2013 0.49 0.02 0.29

The observed hourly concentrations at the Robben Islandand Hangklip sites are presented in Figure **??**, together with the hourly measurements at Cape Point and the measured daily temperature at this site. There is no clear correlation between the peaks in and daily temperature. From March 2012 until June 2013, the mean concentration observed at Robben Island was 391.3 (± 5.02), usually ranging between 389.5 and 394.2 , with a minimum of 382.4 and a maximum of 445.0 . The measurements at Hangklip had a similar mean of 390.6 (± 3.89), usually ranging between 389.5 and 391.4 with a minimum

of 380.4 and a maximum of 430.6 . The Cape Point measurements have a narrower range of 382.9 to 412.3 , with a mean of 392.1 , indicating less influence from local sources and sinks.

Observed hourly concentrations () (left-side axis) at the Robben Island (top closed red circles) and Hangklip (bottom closed black circles) measurement sites. The blue line appearing at the bottom of each plot is the concentration measurements at Cape Point station () and the green line at the top of each plot is the mean daily temperature (°C) as measured at the Cape Point station, which is represented by the right-side axis.

The mean diurnal cycle for each month is presented in Figure **??**. Across all months, the diurnal cycle of concentrations at Cape Point are relatively flat compared to Robben Island and Hangklip. In November 2012 and February 2013 the diurnal cycle for both measurement sites was the most flat. This is the summer period , when temperatures are high and the Western Cape experiences the lowest amount of rain. The amplitude of the diurnal cycle at both sites increased from April, reaching a maximum amplitude in June and July. This is during the winter rainfall period in the Western Cape. Temperatures are mild and much of the vegetation growth occurs during this period. The diurnal cycle of the Hangklip site dipped below both Cape Point and Robben Island, indicating that this site is more affected by local sinks of . Robben Island consistently had the highest peaks in concentrations across all months, indicating that this site was the most affected by local sources, which is what we expected.

Diurnal cycle of the observed concentrations () for each month and at each site with standard error bars, where the standard error is calculated over all measurements available for that hour of the day during that particular month. Cape Point is the generally flat diurnal cycle in blue, Robben Island with the generally larger daytime concentrations in red, and Hangklip with the generally lower afternoon concentrations in black.

The mean diurnal cycle over the whole measurement period is presented in Figure **??**, separated by site and by working week and weekend. The background site, Cape Point, shows no discernible difference between the mean concentrations over the week and weekend, whereas Robben Island and Hangklip sites measure concentrations during the working week which tend to be larger across most of the day. Both the early morning and afternoon means show a clear tendency for these sites to have larger concentrations during the working week compared to the same time of day over the weekend, which can only be due to anthropogenic influences. This supports the separation of fossil fuel fluxes into working week and weekend contributions.

Diurnal cycle of the observed concentrations () over the full measurement period from March 2012 until June 2013 at each site with standard error bars, where the standard error is calculated over all measurements available for that hour of the day during the entire measurement period, separated by site (Cape Point - blue and light blue, Robben Island - red and dark red, Hangklip - black and grey) and by working week (brighter colour) and weekend (duller colour).

Table **??** provides the measured concentration availability at each site. Robben Island had slightly higher data availability compared to Hangklip, with a 65.6% data availability overall for the sixteen month measurement campaign. Each site was equipped with a Hauwei USB modem connected to a 3G network, and were set to submit data to an email address on a regular basis. Through these emails or through connecting remotely to the instruments, instrumentation problems could be detected. Most of the down-time at the sites was attributed to either pump failure, or occasionally the instrument software had failed and there was no available person to restart the instrument due to limited access to these sites. Robben Island lighthouse is

manned, and therefore it was possible to request the lighthouse keeper to restart the instrument when frozen, but more regular than expected pump problems required visits to the site to replace the offending device. This entailed making arrangements for unplanned voyages to Robben Island, which could take some time to arrange. Hangklip is unmanned, and the site has strict access control, therefore problems at this site tended to take slightly longer to remedy. The final four months considered in this study had the best continuity of data availability.

The number of days with available measurement data for each month (out of a possible four week period considered) and overall percentage available data out of 16 four-week periods for each site. Overall Site Percentage 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 Robben Island 28 21 28 14 14 28 28 0 14 0 0 21 28 28 28 28 68.75 Hangklip 28 21 28 14 14 0 28 0 14 0 0 21 28 28 28 28 62.50 Combined 56 42 56 28 28 28 56 0 28 0 0 42 56 56 56 56 65.63

3.1 Reference Inversion

The inversion configuration considered as the reference inversion in this study consisted of the most comprehensive inclusion of information available. Here full monthly inversions were performed solving for four separate weekly fluxes, decomposed into working week and weekend fossil fuel contributions, full week NEE contributions, each separated into day and night contributions. The observation error covariance matrix consisted of variance elements which have been estimated based on wind speed , whether day or night, and on the concentration measurement variance, and temporal correlation, with a one hour correlation length, was included between measurements from the same site. The variance elements of the prior flux covariance matrix were estimated from the fossil fuel inventory analysis and net ecosystem productivity (NEP) estimates from CABLE, with correlation between NEE estimates from adjacent pixelsCape Point over the 16 month period are provided in supplementary material section 1.1. Information is also provided on the assessment of the goodness-of-fit of the prescribed covariance structures in supplementary material section 1.2 to justify the use of the scaling factor of 2 to increase the original estimates of the uncertainties in the prior fluxes to get the χ^2 statistic closer to one. The average wind speed and direction, supplied as monthly maps of the wind fields across the domain, as modelled by CCAM are provided in section 1.3.

3.0.1 Modelled Concentrations

3.1 Modelled Concentrations

The time series of the prior and posterior modelled concentrations at Robben Island were compared to with the observed concentrations and the prior modelled concentrations (Figures 6a and b). The prior estimates tended to be in the correct range for CO_2 concentration measurements, but could be higher or lower compared to with the observations by as much as 100 ppm. The It is possible to test whether our assumed uncertainties in the prior fluxes are consistent with the misfit between the prior modelled concentrations and observations. Michalak et al. (2005) pointed out that the covariances of the differences between the prior simulation and observations is given by the matrix $\mathbf{HC_{s_o}H}^T + \mathbf{C_c}$. This matrix accounts for both the uncertainty in the prior fluxes and in the observations. The square root of the diagonal elements of this matrix had a similar distribution to the absolute mismatches between the observations and prior modelled concentrations, showing that the set-up is statistically

consistent. The prior concentrations tended to spike at the same time as the observations, but these spikes were usually larger in the prior modelled concentrations.

As one would expect, the agreement between the posterior ~~and observed concentrations~~ modelled concentrations and the observations was much stronger compared ~~to~~ with the prior estimates~~, with posterior estimates showing the inversion's ability to track~~. The posterior concentrations appeared to track the observed concentrations during localised ~~events.~~ "pollution" events. For example, in March to April 2012 all except one of the spikes in the observed CO_2 concentration was replicated in the posterior concentrations. The agreement can be assessed by means of the intraclass correlation coefficient (ICC) (Shrout and Fleiss, 1979), which is a stronger condition than correlation. Values close to zero indicate poor agreement while values close to one indicate strong agreement. The ICC ~~is~~ was low at 0.03 (95% CI: 0.01 to 0.06), but still significant, between the observed and prior modelled concentrations, but ~~goes~~ went up to 0.59 (95% CI: 0.57 to 0.61) between the observed and posterior modelled concentrations.

We define prior residuals as the difference between the observed and prior modelled concentrations, and posterior residuals as the difference between the observed and posterior modelled concentrations. A time series plot of the prior and posterior residuals ~~is given in Figures 6 c and d~~given in Figure 6 indicates more clearly how large the misfits between the modelled and observed concentrations can get. The prior residuals could be ~~very~~ large in either the positive or negative direction, up to 100 ppm and occasionally out by as much as 200 ppm. The posterior residuals were much closer to the zero line, with the highest deviation equal to 33 ppm. ~~Summary statistics of the residuals indicate that the mean~~ The bias in the prior ~~estimates~~ modelled concentrations was -2.9 ppm~~with standard deviation equal to~~ . The standard deviation of the prior residuals was 21.4 ppm ~~and~~ (interquartile range between -9.1 and 3.7 ppm~~. From the residualsof the posterior estimates, the bias was reduced~~), indicating a large amount of spread in the residuals. The bias in the posterior modelled concentrations went down to 0.5 ppm ~~,~~ and the standard deviation of residuals reduced to 3.9 ppm ~~, and interquartile range to between~~ (interquartile range -1.5 and 1.5 ppm), showing a significant ~~improvement over the prior estimates. The~~ reduction in the misfit compared with the prior modelled concentrations. Compared with the standard deviation of the observed concentrations, which was 5.02 ppm, the standard deviation of the posterior residuals was lower by 1.1 ppm~~compared with that of observed concentrations , which was 5.02~~ , indicating that the uncertainty in the posterior estimates of the concentrations was well below the expected variability around the observed concentrations.

The time series of the observed, prior and posterior concentrations at Hangklip reveal a similar result compared ~~to~~ with those for Robben Island (~~Figures 7a and b~~Figure 7). The prior estimates ~~can~~ could be much larger or smaller compared ~~to~~ with the observed concentrations~~, but the~~ . The posterior concentration estimates ~~match~~ matched much more closely with the observed concentrations ~~, also showing ability to track local events at the site, even more closely compared to the Robben Islandposterior concentrations~~compared with those for Robben Island. The ICC between the observed and prior modelled concentrations was similar to Robben Island at 0.03 (95% CI: 0.003 to 0.05), but ~~shows much better~~ the agreement between the observed and posterior modelled concentrations was better with an ICC of 0.76 (95% CI: 0.75 to 0.77). The prior residuals at the Hangklip site tended to be less extreme compared ~~to~~ with those for Robben Island, with a maximum deviation of 117 ppm in either direction (~~Figures 7c and d~~Figure 7). The summary statistics of the residuals indicate that the mean bias in the prior estimates

was 2.4 ppm with standard deviation equal to 17.6 ~~and~~ (interquartile range between -2.3 and 6.5 ppm). For the posterior residuals, the bias was reduced to 0.04 ppm with standard deviation equal to 2.46 (interquartile range -1.1 to 0.8), the standard deviation lower by 1.4 ppm compared with the standard deviation of the observed concentrations, which was 3.89 ppm.

The observed and modelled concentrations and their misfits are provided separately for day and night concentrations in Figures 6 and 7. There is no notable difference in the degree of misfit between day and night at either site. The large improvement in the representativeness of the posterior concentrations in relation to the observed concentrations at both sites lends confidence to the reference inversion's ability to ~~improve~~ adjust the estimates of the ~~sources~~ fluxes to better match ~~observed conditions~~ the true fluxes in the region.

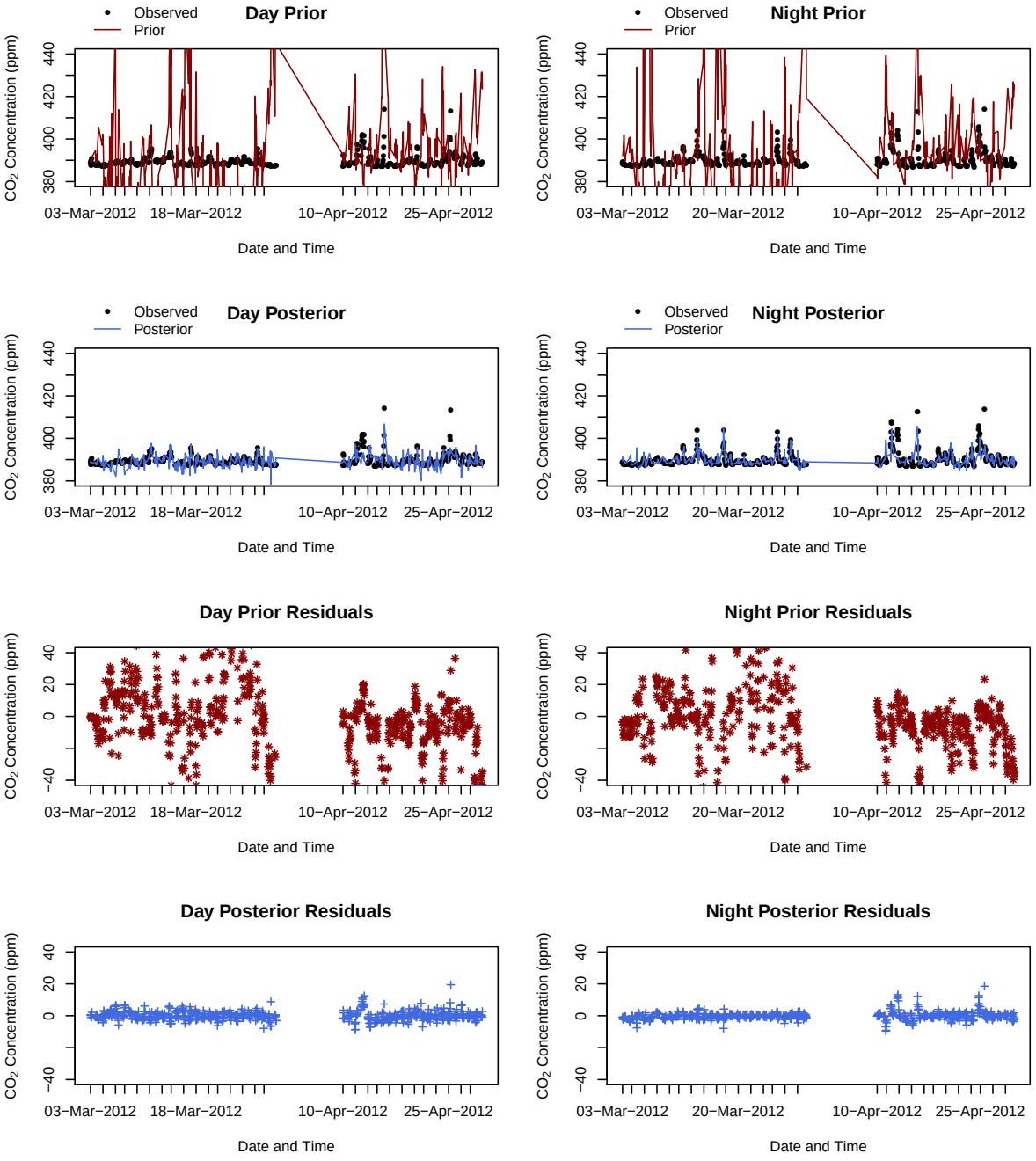


Figure 6. ~~A~~ The top 4 panels provide a time series ~~over the sixteen month measurement campaign~~ of ~~(a)~~ the observed~~and prior modelled concentrations () from the reference inversion at Robben Island~~, ~~(b) the observed~~ prior and posterior modelled concentrations at the Robben Island ~~, (c) the prior~~ site. The time series is separated into day and night-time periods. The residuals ~~(), defined as the difference~~ between the observed and prior/posterior modelled concentrations, ~~and (d) the posterior residuals,~~ defined as the difference between the observed and ~~posterior~~ modelled concentrations, are provided in the lower panel 4 panels. ~~Closed black circles - observed concentration; open red triangles - prior concentrations; purple crosses - posterior concentrations; open red inverted triangles - prior residuals; open purple circles - posterior residuals~~The first two months are presented here and remainder of the time series is presented in the supplementary material section 1.4.

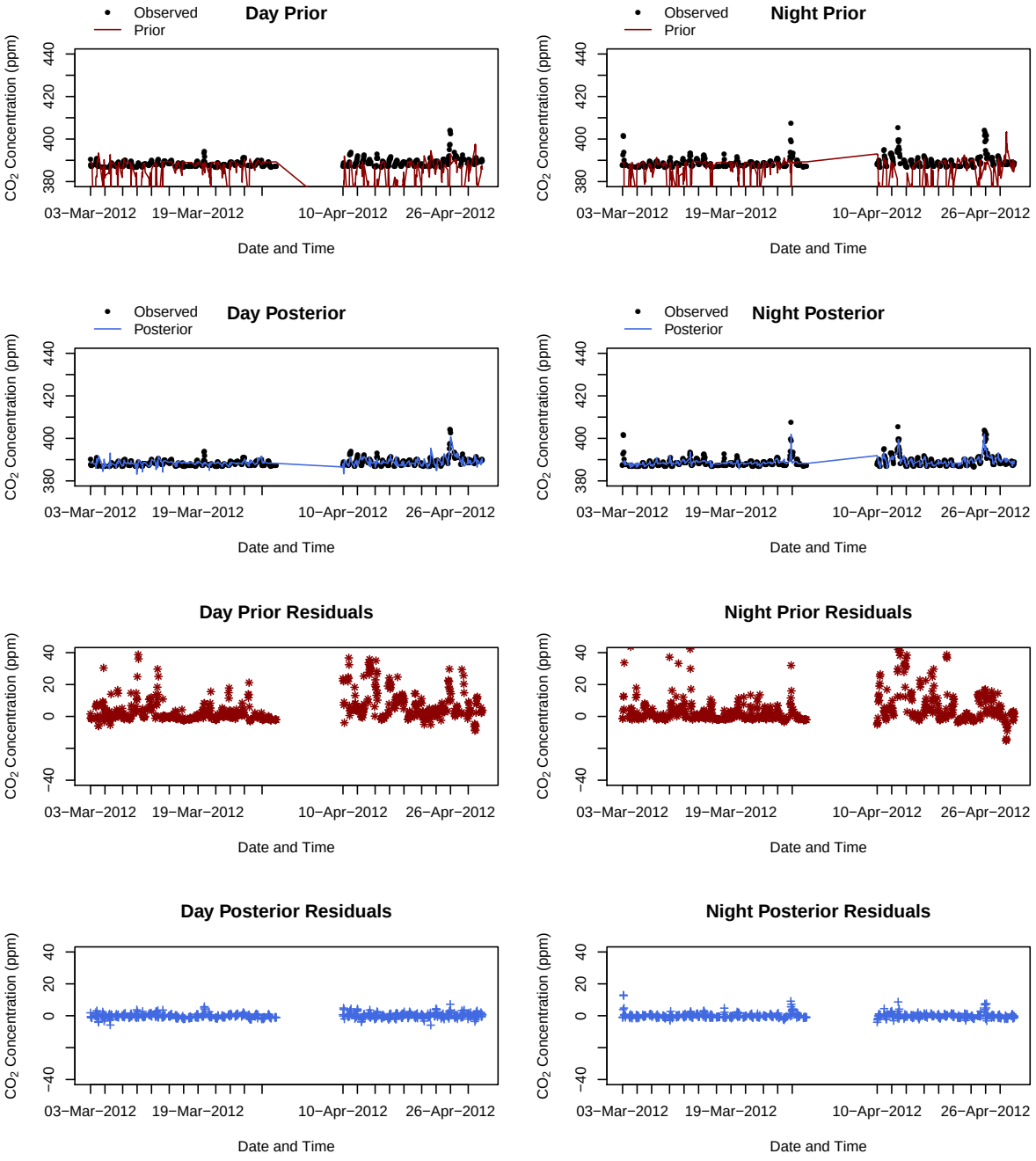


Figure 7. ~~A~~ The top 4 panels provide a time series ~~over the sixteen month measurement campaign~~ of ~~(a)~~ the observed~~and prior modelled concentrations () from the reference inversion at Hangklip~~, ~~(b) the observed~~ prior and posterior modelled concentrations at ~~Hangklip, (c)~~ the ~~prior~~ Hangklip site. The time series is separated into day and night-time periods. The residuals ~~(), defined as the difference~~ between the observed and prior/posterior modelled concentrations, ~~and (d) the posterior residuals,~~ defined as the difference between the observed and ~~posterior~~ modelled concentrations, are provided in the lower panel 4 panels. ~~Closed black circles - observed concentration; open red triangles - prior concentrations; purple crosses - posterior concentrations; open red inverted triangles - prior residuals; open purple circles - posterior residuals.~~ The first two months are presented here and remainder of the time series is presented in the supplementary material section 1.4

The mean working week and weekend diurnal cycles in the observed, prior and posterior modelled concentrations are shown for each site and for each month in the supplementary material (~~Supplementary Material Section 1.2~~supplementary material section 1.5). Figure 8 provides the mean working week and weekend diurnal cycle over the full measurement period. For Robben Island, the mean concentrations for each hour indicate that the emissions are overestimated by the prior estimates~~,~~ ~~therefore producing larger concentrations compared to the observed~~. The posterior modelled concentrations are ~~all~~ much closer to the observed concentrations, ~~reproducing~~ replicating the peak in concentrations to be between 8:00 and 9:00 in the morning and the trough in concentrations to occur between 15:00 and 18:00. Overall the cycle in the posterior concentrations is ~~more flat compared to~~ flatter compared with that of the observed concentrations. The ~~mean observed concentrations for each hour indicates that for most hours of the day, the concentrations~~ observed concentrations during the week are usually slightly higher compared ~~to~~ with those over the weekend. The posterior estimates show a smaller deviation between the week and weekend concentrations at each hour of the day, particularly around mid-morning, compared ~~to~~ with the observed week and weekend concentrations.

The prior estimates for the Hangklip measurement show the opposite bias compared with Robben Island, with prior modelled concentrations lower at each hour compared ~~to~~ with the observed concentrations. The posterior modelled concentrations for Hangklip overlap closely with the observed concentrations. ~~Compared~~ When compared with Robben Island, there is slightly less separation between the working week and weekend concentrations at each hour. This should be expected as the concentrations observed at the Hangklip site are more dominated by biogenic sources compared ~~to~~ with Robben Island. The closest fossil fuel sources are those from transport and domestic emissions. The main road through this area carries a large amount of commercial traffic during the week, and over the weekend the area is frequented by weekend residents and tourists, and therefore anthropogenic activity is not expected to be much lower over the weekend. The posterior concentrations show that the inversion ~~has managed~~ was able to replicate the separation between the mean hourly working week and weekend concentrations shown by the observed concentrations for most hours of the day, particularly the difference between these concentrations occurring between the hours of 7:00 and 9:00.

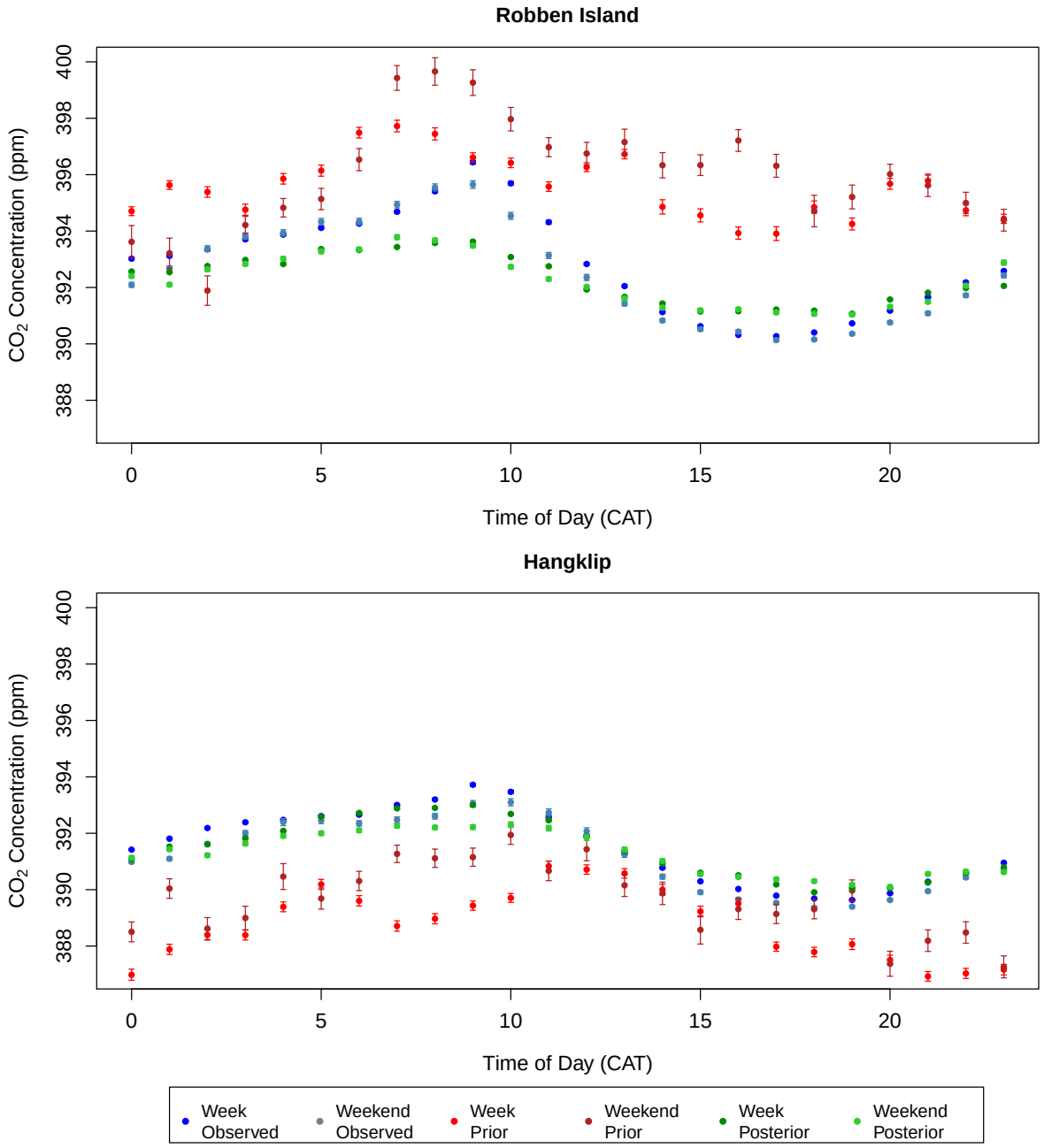


Figure 8. The hourly diurnal cycle (mean ~~estimates~~ concentrations for each hour with ~~standard error bar~~95% confidence interval) in the observed, prior and posterior modelled CO_2 concentrations (ppm) over the full measurement period from March 2012 until June 2013, separated by working week and weekend, and plotted separately for Robben Island (top) and Hangklip (bottom) measurement sites. The diurnal plots are separated into working week and weekend observed concentrations (~~black~~ blue and ~~grey~~light blue), working week and weekend prior modelled concentrations (red and dark red), and working week and weekend posterior modelled concentrations (~~dark~~ green and light ~~purple~~green).

The contributions by the fossil fuel and NEE fluxes to the modelled concentrations were determined. These are displayed in Figures 9 and 10 for Robben Island and Hangklip respectively for March to June 2012. July 2012 to June 2013 are supplied in the supplementary material section 1.6. The prior contributions by these two fluxes were of similar magnitude at both sites, but with the fossil fuel fluxes increasing CO_2 concentrations and the NEE fluxes reducing the concentration. After the inversion, the contributions of the posterior NEE fluxes to the modelled concentrations were much more modified by the inversion compared with those from the posterior fossil fuel fluxes. Moreover, the adjustments made to the NEE fluxes resulted in the contributions to the CO_2 concentrations that were much more similar in magnitude to the fossil fuel fluxes, to the extent that the uptake of CO_2 due to biogenic processes cancelled out the contributions made by the fossil fuel fluxes.

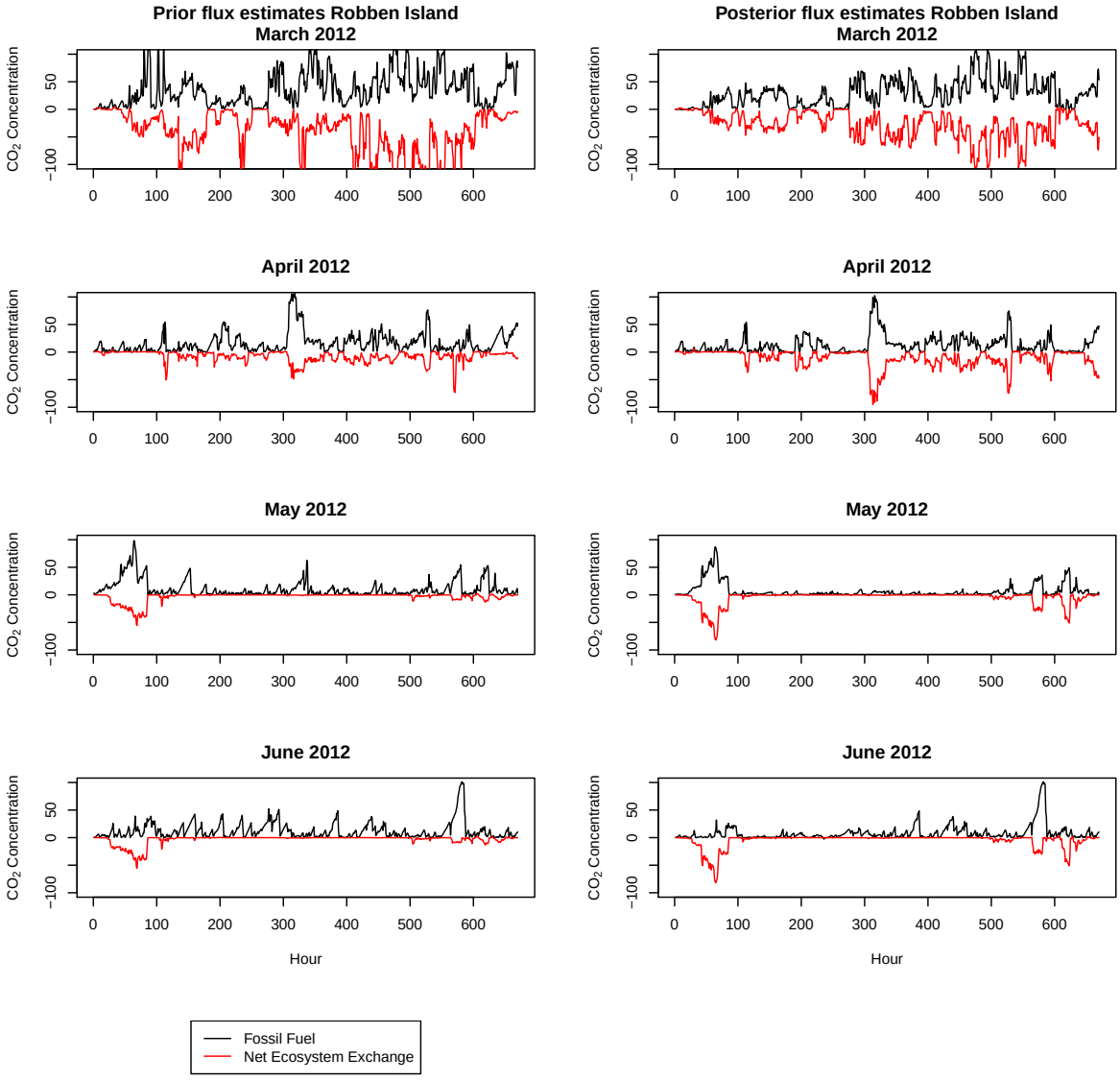


Figure 9. Prior and posterior contributions of the fossil fuel and NEE surface fluxes to the modelled CO_2 concentrations (ppm) at Robben Island from March 2012 until June 2012.

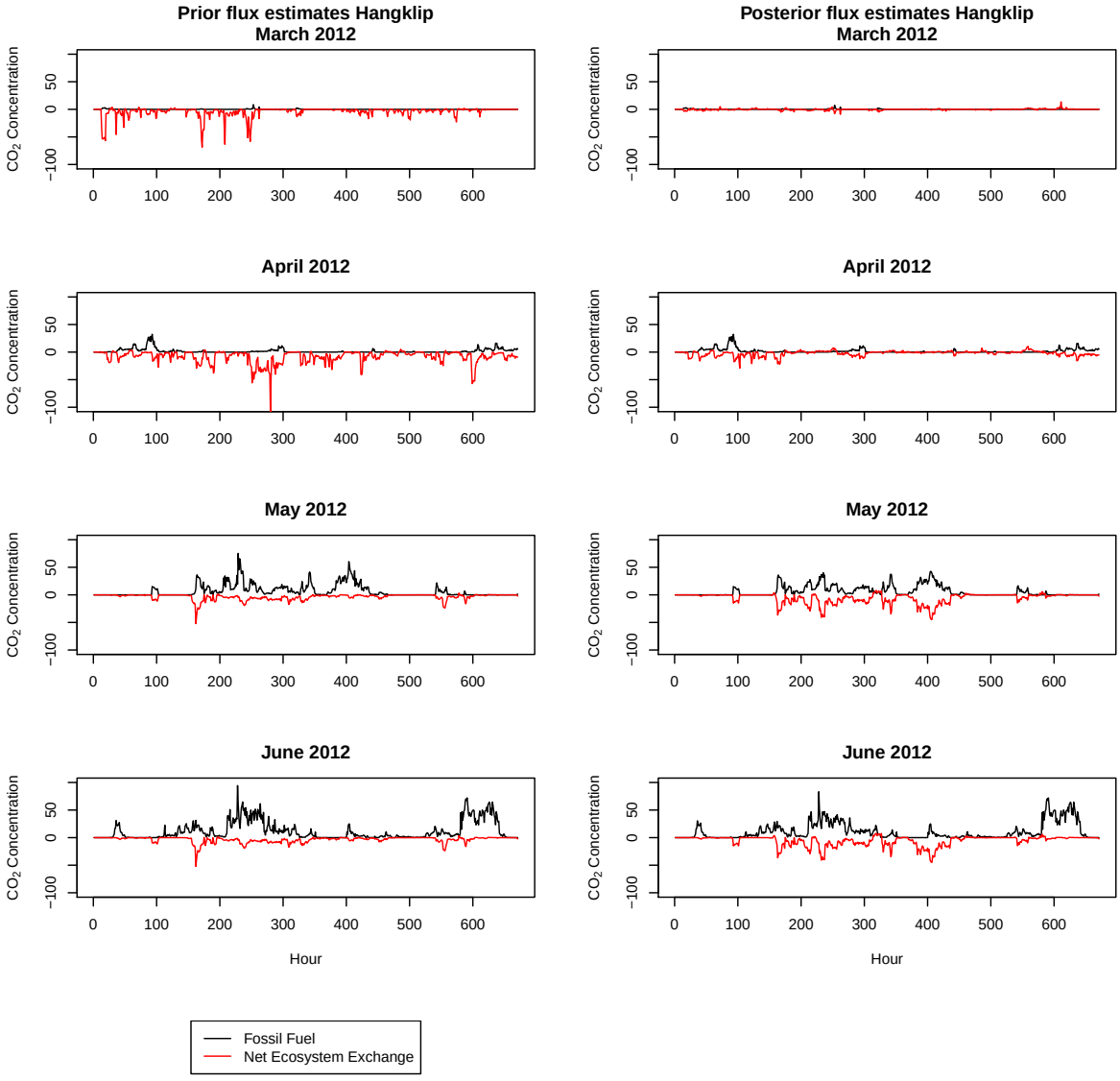


Figure 10. Prior and posterior contributions of the fossil fuel and NEE surface fluxes to the modelled CO_2 concentrations (ppm) at Hangklip from March 2012 until June 2012.

The inversion-corrected ~~background~~ CO_2 concentrations at the domain boundary are obtained from the posterior source estimates. Therefore these concentrations are extracted from the solution of s. Figure 11 provides a time series of the prior and posterior concentration estimates. The ~~north and east border~~ northern and eastern domain borders are terrestrial, whereas the ocean borders the south and west. For all ~~borders~~ four domain boundaries, across the total measurement period, the inversion has ~~generally had very~~ made only small innovations, with the posterior estimates remaining within the 95% confidence limits of the prior concentrations. Only the northern and eastern terrestrial ~~borders show~~ boundaries showed some deviations from the priors between May and June 2012, and between March and April 2013. As these are the terrestrial ~~borders, it would be expected that if the inversion needed to correct for unknown sources~~ boundaries of the domain, the concentrations here would be subject to influences from outside of the domain, ~~it would be at these border locations, whereas we expect the ocean borders to have background concentrations very close to those~~ and it would therefore be expected that the inversion would need to provide greater adjustments to these boundary concentrations. The ocean boundaries would be expected to have concentrations close to the background concentrations provided by the Cape Point ~~background~~ measurement site.

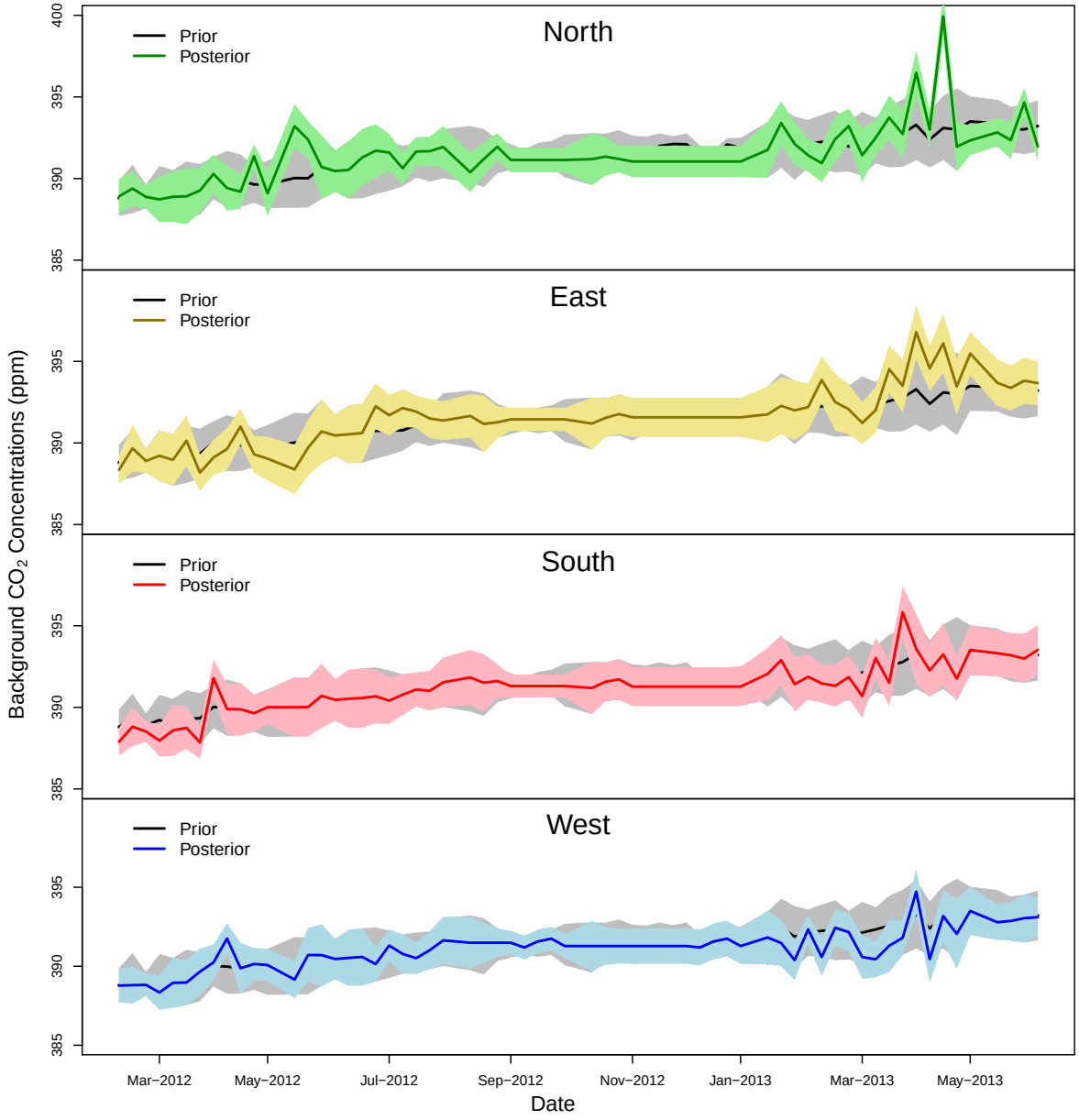


Figure 11. Time series with 95% confidence interval (represented by the shaded area) of the prior (black line) and posterior estimates of the ~~background~~ CO_2 concentrations (ppm) at the domain boundaries (north - green, east - yellow, south - red, west - blue). The prior estimates are the same for each cardinal direction, and are obtained from the Cape Point percentile-filtered observations. The posterior estimates for the concentrations are solved for as additional unknowns in the reference inversion.

3.1.1 ~~Weekly flux Estimates~~

~~The innovation of the inversion can be observed through the differences~~

3.2 Weekly flux Estimates

We refer to the difference between the prior and posterior flux estimates ~~, which we refer to~~ as the innovations~~, and through the change in the uncertainty estimates. Positive innovations indicate that the prior estimates were too large, and that the estimate should be more in the negative direction , whereas negative innovations indicate that prior estimates were set too small and should be more in the positive direction~~. The impact of the inversion on the flux estimates can be assessed through the size and direction of these innovations and through the reduction in the flux uncertainties. Figure 12 shows the ~~changes in~~ innovations in the total flux estimates for each pixel in $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ for the month of May 2012, as well as the percentage ~~change~~ reduction in the flux uncertainty, the percentage ~~change~~ reduction in the fossil fuel emissions and change in NEE fluxes. The mean total weekly flux of a pixel was obtained by ~~summing together all of the flux components of the state vector and dividing by the number of weeks, and the innovations calculated from these mean~~ first deriving the total flux from the six fossil fuel and NEE flux estimates for that week, and then taking the mean of the four weekly fluxes obtained for the month. The innovations for the total flux were calculated as the difference between the total prior and total posterior weekly fluxes. Positive innovations indicate that the prior fluxes were too far in the positive direction, and that the flux should be adjusted towards the negative direction, whereas negative innovations indicate that prior fluxes were too far in the negative direction and should be adjusted towards the positive direction. The uncertainties for the total flux in each pixel are represented by standard deviations which have been derived from the ~~variance estimates for the weekly fluxesfor each pixelover the month of May 2012.~~ elements of the prior and posterior uncertainty covariance matrices of the fluxes. The figure displays the percentage uncertainty reduction within each pixel. The percentage reductions in the total fossil fuel flux are presented, rather than the absolute changes from the prior to posterior flux estimates in each pixel as these changes were small and the inversion generally resulted in a reduction in the fossil fuel flux estimate. The inversion had more freedom to make changes to the NEE estimates as the uncertainties prescribed were large relative to the prior NEE flux estimates, and therefore the absolute changes from prior to posterior are presented in $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$.

May falls within the winter rainfall season of the Western Cape region. The innovations indicate that the total flux for the pixel over the petrol refinery, which had the largest prior flux estimate, was overestimated by the prior ($9.43\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) relative to the posterior estimate ($6.62\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) by an amount of $2.81\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}$ (Fig. 12~~a). Adjustments~~). Innovations were generally small ranging between -0.001 and $0.003\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$. The area around the natural reserves, such as Cape Point and Kogelberg Nature Reserves, had ~~adjustments~~ innovations by the inversion that were close to zero or slightly negative, indicating that the amount of carbon uptake (i.e. negative NEE flux) in these regions was overestimated by the CABLE model. The largest negative ~~adjustment~~ innovation by the inversion was $-0.08\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, from -0.03 up to ~~-0.04~~0.04 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, over a pixel in the Cape Point Nature Reserve. The prior flux indicated the pixel to be a slight carbon sink of $\mathrm{CO_2}$ whereas the inversion adjusted the flux to be a slight $\mathrm{CO_2}$ source. With respect to the rest of the

domain, excluding the crude oil refinery, the ~~largest adjustments to the prior fluxes~~ most substantial innovations were made over the ~~CBD~~ central business district (CBD) area to the south east of Robben Island.

~~In terms of the percentage standard deviation reduction (Fig. 12b),~~ These innovations were positive, indicating that the fossil fuel fluxes were overestimated by the ~~largest reductions occurred over the natural areas, particularly Cape Point to the south of Robben Island, where the posterior uncertainty over the area was significantly lower, by over 50%, compared to the prior uncertainties. Significant reductions are also shown over largely agricultural areas to the north of the CBD region. Over the CBD area itself, the reductions were present, reaching values of close to 60% over a few central CBD pixels, but generally smaller compared to the uncertainty reductions over natural areas in the domain which reached levels as high as 92%.~~ prior estimates.

The ~~sources were split into those which originate from fossil fuels and those from biogenic sources. The~~ percentage change in the fossil fuel estimates, from prior to posterior, indicate the changes tended to be small, except on Robben Island itself, where percentage changes were up to 75% indicating that the emissions on the island were significantly reduced by the inversion (Fig. 12~~e~~). There were a few pixels which had negative change, indicating that the inversion increased fossil fuel emissions, just north of the island. Located on the north eastern shore of Robben Island is a diesel-fuelled power generation plant, as well as desalination plant which is powered by this station. Increases in the fossil fuel fluxes may be due to emissions arising from these activities which have not been accounted for in the inventory analysis. In the inventory analysis there was no fuel information available for any industrial sources on Robben Island, but fossil fuel emissions were included due to domestic and transport ~~sources~~activities, therefore these could have been adjusted by the inversion. The inventory analysis does not take into account explicitly the shipping routes going into CT harbour, or into Robben Island harbour, but rather all the emissions are concentrated within CT harbour, where the shipping information is available. This could also result in the inversion adjusting emissions on the island to deal with near shipping related emissions. There was a region in the Western Cape interior to the east of Robben Island which had slightly increased fossil fuel emissions. The inversion had the effect of mainly reducing fossil fuel emissions along the south east transect extending from Robben Island over the CBD towards Hangklip, or leaving the emission unchanged.

The inversion had a much larger impact on the terrestrial ~~biogenic~~ NEE fluxes (Fig. 12~~d~~). This is unsurprising, as the relative uncertainties of the prior ~~biogenic~~ NEE fluxes were much larger compared with those of the prior fossil fuel ~~estimates~~fluxes. The area of the domain experiencing innovation from the inversion was also much more widespread compared with the innovations made to the fossil fuel ~~emissions~~fluxes. This is in part due to the correlation specified between the ~~biogenic~~ NEE fluxes of adjacent pixels, but not between their fossil fuel fluxes. The majority of the innovation over the domain was close to zero, between -0.02 and 0.02 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, indicating that the inversion was making small absolute ~~changes to the biogenic~~ adjustments to the NEE flux estimates. Over the CBD region, the ~~changes~~ adjustments were the largest, up to 0.32 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ and these differences were positive indicating that the posterior ~~estimates were smaller~~ fluxes were more in the negative direction and therefore the inversion was acting to reduce ~~emissions~~ total emissions of $\mathrm{CO_2}$ over the CBD region relative to the prior estimates, through changes to the ~~biogenic~~ NEE fluxes. The natural region around Cape Point and within the Kogelberg Nature Reserve north of Hangklip showed ~~slightly~~ slight negative changes in the ~~biogenic~~ NEE fluxes,

indicating that the ~~drawdown~~ uptake of CO_2 was reduced by the inversion ~~(i.e. the carbon uptake was less than predicted by CABLE) making the overall~~ making the total flux in these regions more positive towards the atmosphere.

In terms of the percentage standard deviation reduction (Fig. 12), the largest reductions occurred over the natural areas, particularly Cape Point to the south of Robben Island, where the posterior uncertainty over the area was significantly lower, by over 50%, compared with the prior uncertainties. Significant reductions are also shown over largely agricultural areas to the north of the CBD region. Over the CBD area itself, the reductions were present, reaching values of close to 60% over a few central CBD pixels, but generally smaller compared with the uncertainty reductions over natural areas in the domain which reached levels as high as 92%.

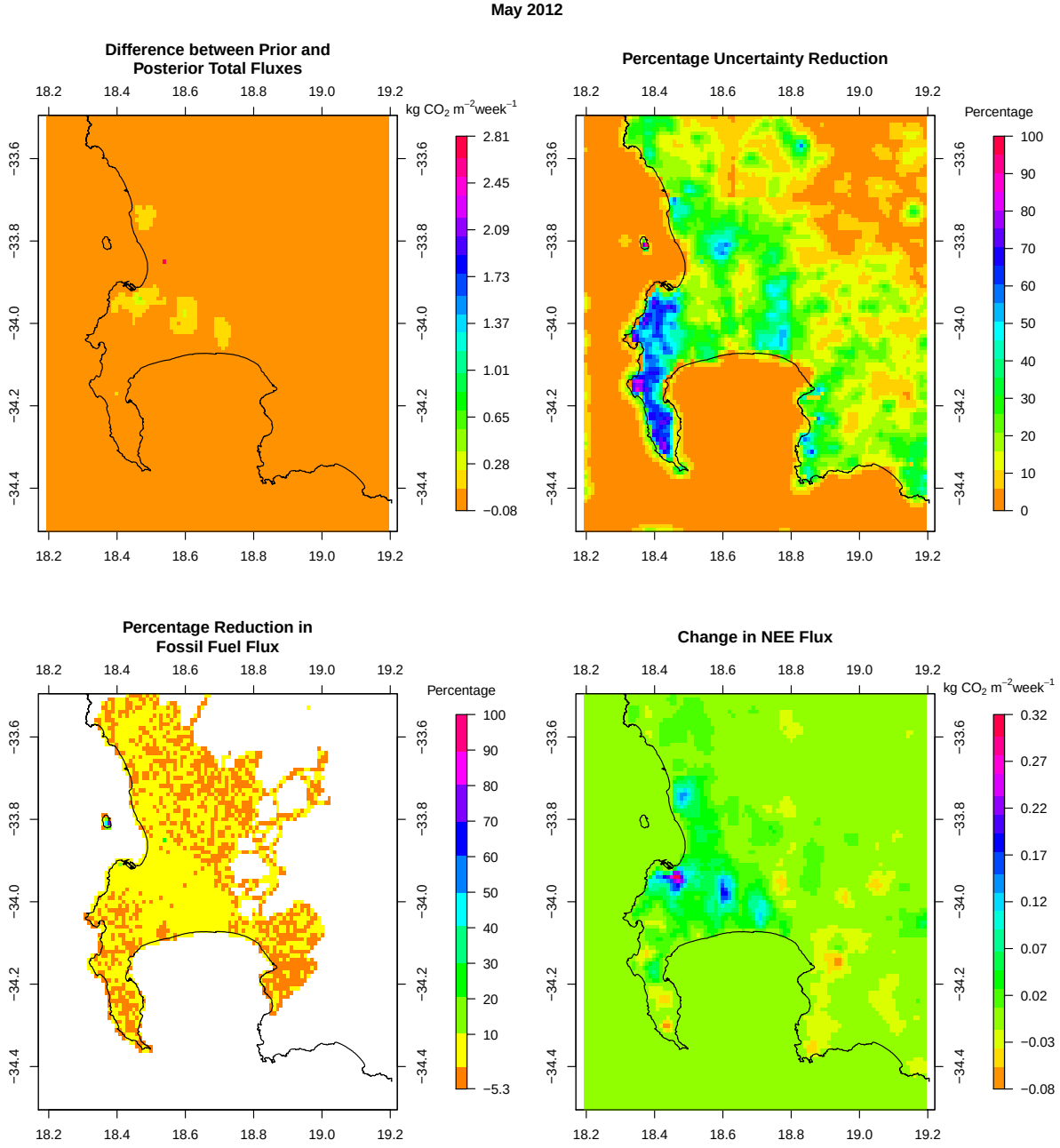


Figure 12. (~~a~~Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for May 2012 (prior - posterior). (~~b~~Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior~~(calculated as (1 - posterior uncertainty / prior uncertainty)×100)~~. (~~e~~Bottom left) Percentage reduction in the fossil fuel ~~emission~~ flux estimates from prior to posterior~~(calculated as (1 - posterior estimate / prior estimate)×100) where estimates were in~~ . (~~d~~Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating ~~larger~~ posterior ~~estimates of~~ CO_2 fluxes ~~relative to~~ were made more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

September is the beginning of the summer months in the Western Cape region, when temperatures start to rise, and the mean monthly rainfall reduces. The innovations across the region during this month were dominated by negative values with the majority ranging between -0.03 and -0.01 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, indicating that the inversion increased emissions of $\mathrm{CO_2}$ over the majority of the domain~~, with a maximum reduction of 2.1 over the crude oil refinery site, from 9.4 down to 7.2 , and a maximum~~ . The maximum increase in $\mathrm{CO_2}$ fluxes of 0.2 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ occurred over an agricultural region north east of the CBD region, from -0.10 up to 0.08 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, where mainly vineyards are located (Fig. 13~~a). There are~~). The inversion increased the total fluxes by an amount close to 0.2 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ over a further five areas~~where the inversion had similar sized increases in the total emissions. One~~ . One of these areas, also over agricultural land, lies further inland to the east of the area with largest increase in ~~emissions~~$\mathrm{CO_2}$ fluxes. Three more regions lie to the north, north east and east of the Hangklip site. These three regions are all within the Hottentots-Holland Mountain catchment area which is largely dominated by vineyard agriculture. The fifth area is located within the Kogelberg Nature Reserve near the Hangklip site. The ~~CBD area of CT had very small reductions in emissions or the posterior emissions remained similar to the priors.~~

~~Reductions in the total pixel standard deviation ranged mainly between 2.3 and 18.6%, with a~~ maximum reduction of ~~88.5%~~ 2.1 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ occurred over the crude oil refinery site, from 9.4 down to 7.2 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ (Fig. 13~~b~~). The ~~largest uncertainty reductions induced by the inversion occurred over the natural areas bordering on the CBD , particularly over the Table Mountain National Park, as well as the natural areas surrounding the Hangklip site. The areas to the east of Robben Island over the Durbanville and Bellville townships, which consists of a mix of residential suburbs, vineyard agricultural areas and industrial areas also showed reductions in the uncertainties of the fluxes~~ inversion made small reductions to the fluxes near the CBD area of CT, or left the posterior fluxes to be similar to the priors.

~~Adjustments~~ Innovations made to the fossil fuel fluxes were mainly made on the transect of the city running between Robben Island and Hangklip, as well as to fossil fuel emissions on Robben Island itself, similar to the month of May. The maximum percentage adjustment to the fossil fuel fluxes was 51.1%, and the mean ~~adjustment~~ innovation close to zero, with almost all ~~adjustments~~ innovations positive, indicating that the posterior estimates were smaller relative to the priors. Robben Island itself showed a ~~mixed~~ mix of positive and negative ~~adjustments~~innovations, with posterior fluxes larger than the priors on the west of the island but smaller than the priors on the east of the island.

The ~~biogenic flux adjustments by the inversion consisted of reductions in emissions~~ effect of the inversion on the NEE fluxes was to make these fluxes more negative over the CBD region, and ~~increases in~~ made $\mathrm{CO_2}$ ~~emissions~~ fluxes more positive relative to the ~~prior estimates~~ priors over the Table Mountain region, over the agricultural area to the north of the CBD, and over the natural regions near Hangklip~~and generally to the eastern region of the domain~~. The fluxes over the CBD were ~~decreased~~ made more negative by up to 0.21 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ through adjustments to the ~~biogenic~~ NEE fluxes. The areas to the north of the domain which had prior NEE estimates between -0.20 and -0.10 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ were made more positive, with posterior estimates ranging between -0.10 and 0.10 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, where the largest changes were over the vineyard agricultural areas. The natural area surrounding Hangklip site showed negative ~~differences~~ innovations between the prior and posterior ~~estimates~~fluxes, indicating that the inversion ~~corrected~~ adjusted the negative prior ~~biogenic~~ NEE fluxes by making

these more positive~~, and so~~ . Therefore the CABLE model appears to have overestimated the amount of CO_2 ~~drawdown~~ uptake during this period.

Most reductions in the total flux uncertainty ranged between 2.3 and 18.6%, with a maximum reduction of 88.5% (Fig. 13). The largest uncertainty reductions induced by the inversion occurred over the natural areas bordering on the CBD, particularly over the Table Mountain National Park, as well as the natural areas surrounding the Hangklip site. The areas to the east of Robben Island over the Durbanville and Bellville townships, comprised of a mix of residential suburbs, vineyard agricultural areas and industrial areas, also showed reductions in the uncertainties of the fluxes.

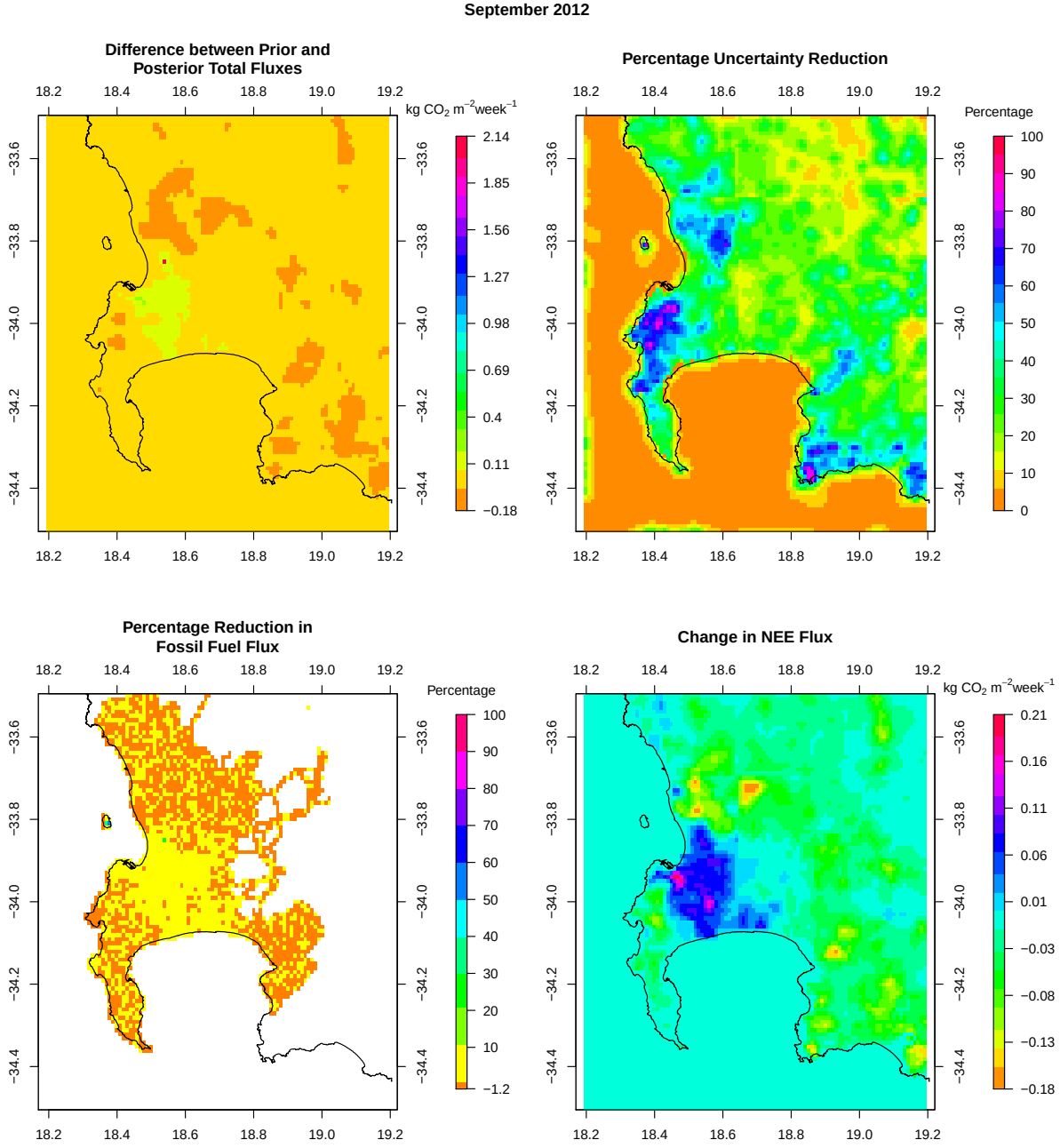


Figure 13. (~~a~~Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for September 2012 (prior - posterior). (~~b~~Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior~~(calculated as (1 - posterior uncertainty / prior uncertainty)×100)~~. (~~c~~Bottom left) Percentage reduction in the fossil fuel ~~emission~~ flux estimates from prior to posterior~~(calculated as (1 - posterior estimate / prior estimate)×100) where estimates were in~~ . (~~d~~Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating ~~larger~~ posterior ~~estimates of~~ CO_2 fluxes ~~relative to~~ were made more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

~~Table ?? presents summary statistics~~ Figure 14 presents boxplots of the pixel-level weekly fluxes over the domain for each month, and the supplementary material (~~Supplementary Material Section 1.3~~supplementary material section 1.7) displays the spatial extent of the adjustments made to prior flux estimates by the inversion, as well as the uncertainty reductions, for each month. The limits of the range for the posterior pixel-level weekly fluxes are more negative compared with those of the prior estimates, indicating that for all the months, the inversion tended to reduce the emission of CO_2~~, and~~. This is also evident from the maps of ~~change~~ differences in total flux estimates between the prior and posterior estimates. Specifically, the inversion tended to reduce the fossil fuel emissions, evident from the ~~smaller mean flux value over the domain~~shift downwards in the distribution of the posterior pixel-level fossil fuel fluxes compared with the prior fluxes. The variability ~~between pixels was also reduced~~ in the posterior fluxes across pixels was slightly reduced compared with the variability in the prior fluxes. The pixel-level ~~natural~~ NEE fluxes, although generally smaller in magnitude and range, were the most altered by the inversion, where these changes were related to the season in which the month fell. For most months, the ~~biogenic~~ overall mean in the NEE fluxes became more positive, indicating less uptake of CO_2 than predicted by the CABLE model. But for the months of June and July, which occur mid-winter, the mean NEE was made more negative. The minimum values were also at least twice the minimum value of the prior estimates for all months, and this is also evident from the maps of change in NEE from prior to posterior, which show that the inversion ~~is reducing~~ reduced the CO_2 flux towards the atmosphere over the CBD region by altering the ~~biogenic~~ NEE fluxes within those pixels.

For most months the percentage uncertainty reduction for the pixel-level total flux reached over 90% for at least some of the pixels, with a maximum uncertainty reduction over a pixel of 97.7% in March 2012. The lowest maximum reduction in the uncertainty of the total flux in a pixel in a month was 68.2% in June 2012.

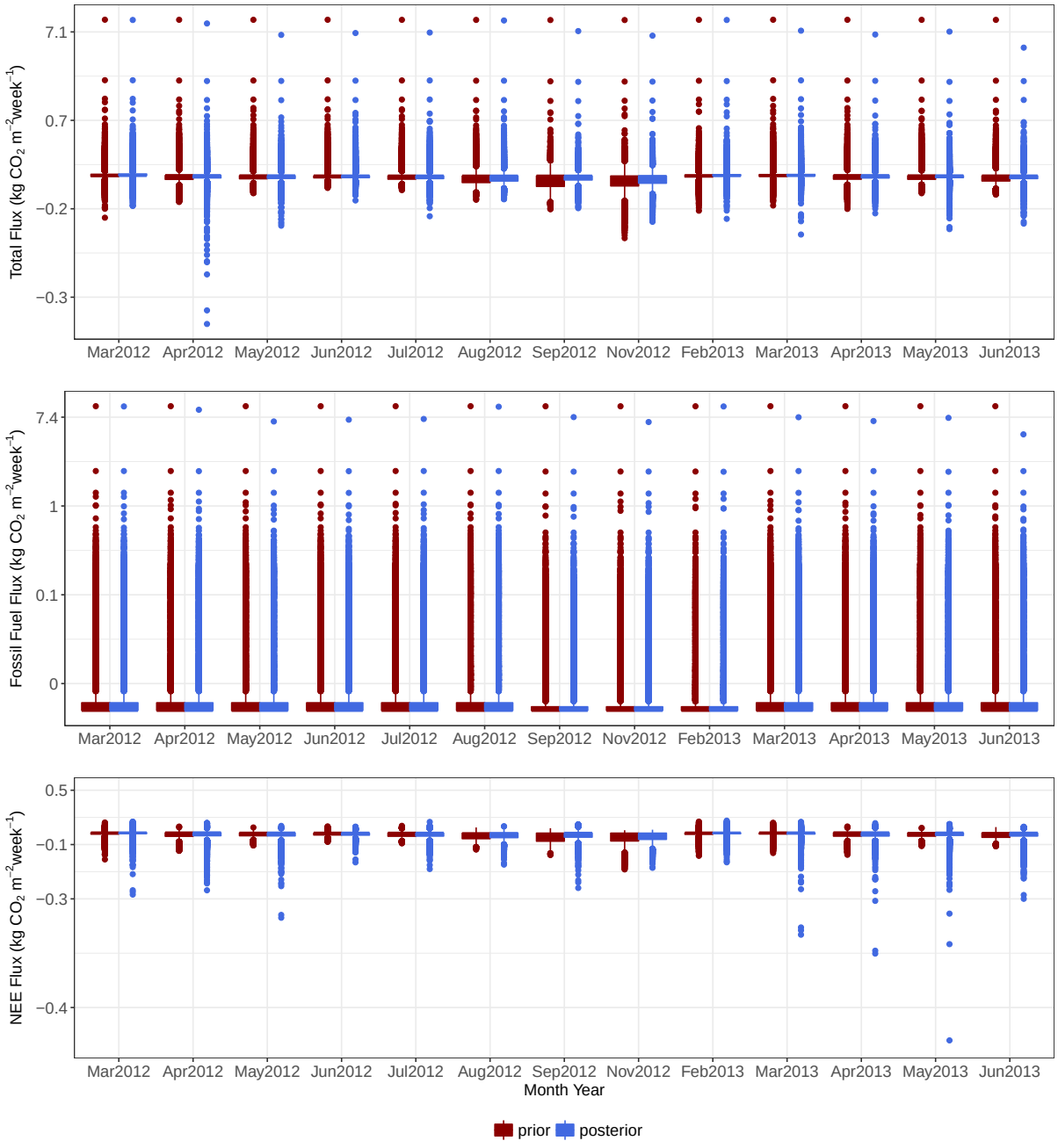


Figure 14. ~~Summary statistics~~ Boxplots of the pixel-level weekly prior and posterior flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for each month. These ~~are~~ plots present summary statistics calculated over all pixels in the domain of the ~~average total flux within~~ pixel-level mean weekly fluxes. The y-axis is presented on a ~~pixel~~log scale.

~~Month Prior Posterior Prior Posterior Prior Posterior mean (sd) mean (sd) mean (sd) mean (sd) mean (sd) mean (sd) (min; max) (min; max) (min; max) (min; max) (min; max) (min; max) Mar2012 0.0077 (0.1123) 0.0102 (0.1059) 0.0144 (0.1083) 0.0139 (0.1053) -0.0066 (0.0248) -0.0037 (0.0342) (-0.1755; 9.4523) (-0.2089; 9.2106) (0; 9.4510) (0; 9.2408) (-0.1762; 0.0958) (-0.4411; 0.2984) Apr2012 -0.0004 (0.1114) 0.0015 (0.0879) 0.0143 (0.1081) 0.0137 (0.0829) -0.0147 (0.0249) -0.0122 (0.0526) (-0.1283; 9.4345) (-0.4673; 6.3828) (0; 9.4510) (0; 6.5232) (-0.1283; 0.056) (-0.4673; 0.2463) May2012 0.0021 (0.1094) 0.0011 (0.0536) 0.0143 (0.108) 0.0132 (0.0557) -0.0123 (0.0174) -0.0121 (0.0435) (-0.0932; 9.4318) (-0.3663; 2.0896) (0; 9.4510) (0; 2.1800) (-0.0932; 0.0479) (-0.5605; 0.1665) Jun2012 0.0073 (0.1092)~~

An example pixel, located near the CBD sources, was selected in order to investigate posterior ~~flux covariances estimated by the~~ uncertainty spatial covariances in the fluxes resulting from the inversion. For a given week, ~~each pixel~~ the total CO_2 flux is composed of ~~potentially~~ six sources: working week and weekend fossil fuel sources, both day and night, and the day and night biogenic sources; and each of these sources could have a non-zero covariance term between itself and the same source but from surrounding pixels, or with one of the other five sources from the same pixel or from surrounding pixels. ~~The posterior covariance terms are determined by the product of the sensitivity matrix and observation error covariance matrix, and by the prior flux covariance matrix. Once the covariances are available, the correlations can be calculated between the reference source and all other sources. Strong correlations on uncertainties indicate that the two parameters have only been resolved as some linear combination of the two (Tarantola, 2005).~~

The posterior uncertainty covariances between the ~~day~~ daytime working week flux of the ~~reference~~ selected pixel and other sources are only notably different from zero for working week daytime fossil fuel fluxes, working week night-time fossil fuel fluxes, and the daytime biogenic fluxes. The uncertainty covariances between the daytime working week fluxes reveal that non-zero covariances do not necessarily have to be close in proximity to the ~~reference~~ selected pixel, and negative and positive ~~covariance pixels can cluster around each other. The~~ covariances can cluster in space. The uncertainty covariances with the night-time fossil fuel fluxes were larger than those during the day, but were limited to a few pixels ~~around the reference~~ close to the selected pixel. These covariances ranged between -0.15 and 0.09 $g^2\ CO_2\ m^{-4}\ week^{-2}$. The non-zero ~~covariances with biogenic sources~~ uncertainty covariances with NEE fluxes were larger (between -1.50 and 0.88 $g^2\ CO_2\ m^{-4}\ week^{-2}$) and fluctuated between patches of positive and negative values. ~~Close~~ Closer to the CBD area there was a ~~distinctive~~ distinct region of positive uncertainty covariance between the fossil fuel source of the ~~reference~~ selected pixel and the ~~biogenic~~ NEE fluxes from a region over the Table Mountain area and a negative covariance patch south of the CBD. The eastern terrestrial part of the domain had patches of positive and negative covariances~~between the fossil fuel source of the reference CBD pixel and the daytime biogenic fluxes. Considering the correlations, these are very smallwhen comparing pixel to pixel correlations~~. When converting these covariances are into correlations, correlations are small; no bigger than 0.001 in either direction.

The covariances between fossil fuel and NEE flux uncertainties are small because the uncertainties in the prior modelled concentrations that are attributed to the flux contributions ($\mathbf{HC_{s_o}H}^T$) are small relative to the uncertainties specified for the modelled concentration errors ($\mathbf{C_c}$). This is not because our prior uncertainty is small but because the transport Jacobian only projects fluxes from individual pixels weakly into modelled concentrations. As the uncertainty in the modelled concentration errors is decreased, the size of the posterior off-diagonal covariance elements between the fossil fuel and NEE flux uncertainties from the same pixel increases. This can easily be confirmed through the use of a toy inversion system using typical values for $\mathbf{H}$, $\mathbf{C_{s_o}}$ and $\mathbf{C_c}$ from our inversion system. This is presented in section 1.8 of the supplementary material.

The sum of the covariances ~~between the reference~~ in the uncertainties between the selected fossil fuel flux and all other ~~sources~~ fluxes equals -25.8 $g^2\ CO_2\ m^{-4}\ week^{-2}$. Therefore the covariances ~~for this pixel~~associated with this pixel's fossil fuel flux would reduce the ~~total variance for the total daytime fossil fuel source~~ variance associated with the total pixel-level flux by 51.7, where the total ~~variance for~~ uncertainty of the fossil fuel ~~source of this pixel~~ flux, as a variance, was 233.7 $g^2\ CO_2\ m^{-4}\ week^{-2}$. ~~As the prior covariance matrix had no non-zero covariance terms between any of the fossil fuel sources,~~

these covariances in the posterior covariance matrix for this source are all as a consequence of the projection of the observation error covariance matrix into the flux uncertainty space by the sensitivity matrix in the term $\mathbf{H}^T\mathbf{C_c}^{-1}\mathbf{H}$ of the solution for $\mathbf{C_s}$. The non-zero covariance terms indicate which of the sources could not be solved independently by the inversion, but where the inversion solved for a linear combination of these sources.

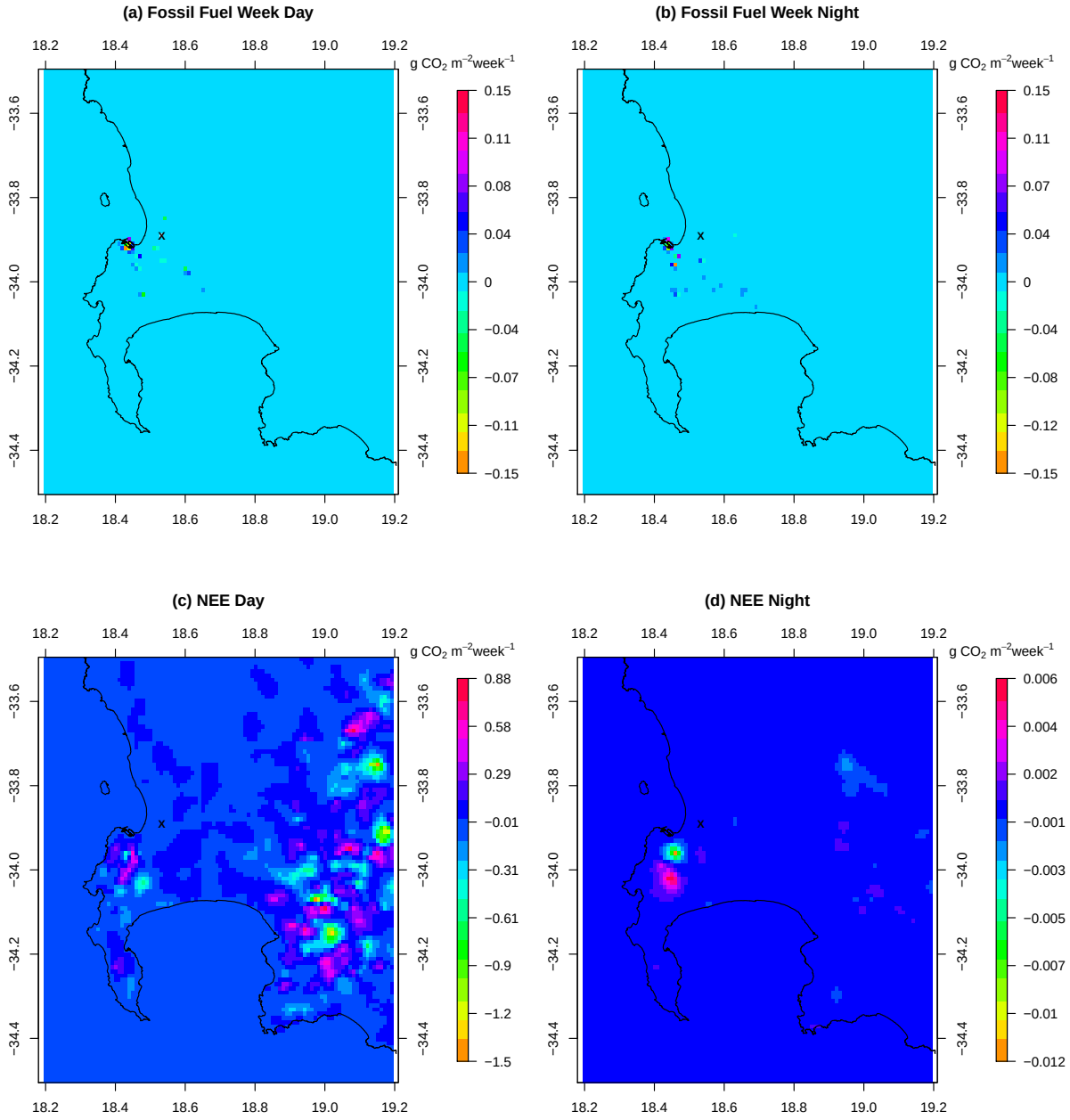


Figure 15. ~~A single pixel was selected near the Cape Town CBD during the first week in March 2012. The posterior~~ Posterior covariances ~~were extracted for this pixel, and the covariances considered between the daytime working week fossil fuel flux within this pixel and all other sources~~ (~~i.e. the remaining sources in this pixel, and all sources within all other pixels). The three plots are (a) covariances (~~g^2 CO_2 m^{-4} $week^{-2}$) between the uncertainty in the fossil fuel working week daytime ~~source~~ flux during the first week of March 2012 in the selected pixel (marked by X) and ~~surrounding~~ (a) all other fossil fuel working week daytime ~~sources from other pixels~~fluxes; (b) ~~covariances between~~ fossil fuel working week ~~daytime source in the selected pixel and fossil fuel working week~~ night-time ~~sources~~ fluxes within this pixel and all other pixels; (c) ~~covariances between fossil fuel working~~ week NEE daytime ~~source in the selected~~ fluxes within this pixel and ~~biogenic flux daytime sources~~ all other pixels; (d) NEE night-time fluxes within this pixel and all other pixels. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

~~A single pixel was selected near the Cape Town CBD during the first week in March 2012. The posterior correlations were extracted for this pixel, and the correlations considered between the daytime working week fossil fuel flux within this pixel and all other sources (i.e. the remaining sources in this pixel, and all sources within all other pixels). The three plots are (a) correlations between fossil fuel working week daytime source during the first week of March 2012 in the selected pixel and surrounding fossil fuel working week daytime sources from other pixels; (b) correlations between fossil fuel working week daytime source in the selected pixel and fossil fuel working week night-time sources within this pixel and all other pixels; (c) correlations between fossil fuel working week daytime source in the selected pixel and biogenic flux daytime sources within this pixel and all other pixels. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.~~

3.2.1 ~~Weekly Totals~~

3.3 Weekly Totals

Three time series plots of the total weekly estimates of CO_2 flux over the full ~~domain~~ spatial domain (i.e. all fluxes aggregated over all pixels) are presented in Figure **??**. The total flux estimate for a week represents the sum of all the fossil fuel and ~~biogenic fluxes for all periods during~~ NEE fluxes for that week. ~~The uncertainties for these total estimates accounts for correlations between fossil fuel and biogenic fluxes.~~ For the prior ~~estimates the only correlations~~ total flux uncertainties, the only non-zero covariances are those imposed between ~~biogenic~~ the uncertainties in NEE fluxes of neighbouring pixels. ~~The variance for each total estimate accounts for those non-zero covariance terms~~ These positive covariance terms increase the total uncertainty. The uncertainties in the posterior ~~covariance matrix between the different flux components contributing to the total~~ total fluxes includes the posterior covariance terms between the uncertainties in the fossil fuel and NEE fluxes, which are generally negative and have the overall effect when summed of reducing the uncertainty in the total flux estimate.

Uncertainties in the prior total fossil fuel fluxes ranged between 1.3 and 1.5 $\mathrm{kt\ CO_2\ week^{-1}}$, whereas the posterior uncertainties ranged between 0.9 and 1.5 $\mathrm{kt\ CO_2\ week^{-1}}$. For the NEE fluxes the uncertainties ranged between 23.6 and 57.3 $\mathrm{kt\ CO_2\ week^{-1}}$ before the inversion and between 15.8 and 47.1 $\mathrm{kt\ CO_2\ week^{-1}}$ after the inversion. Uncertainties in the total fluxes were similar to those of the NEE fluxes, with prior uncertainties ranging between 23.6 and 57.3 and posterior uncertainties ranging between 15.8 and 47.1 $\mathrm{kt\ CO_2\ week^{-1}}$. The median percentage uncertainty reduction in the total weekly flux was 28.0 % and ranged between 2.3 and 50.5 %, with the largest reduction occurring in March 2012.

The posterior total estimate for the emission of CO_2 from the domain ~~is~~ was within the confidence bounds of the prior total estimate for the majority of the period from March 2012 until June 2013. The confidence bounds of the posterior total estimates ~~are narrower compared to~~ were narrower compared with those of the prior total estimates. Total prior ~~estimates range~~ flux estimates ranged between 139.5 and -386.8 $\mathrm{kt\ CO_2\ week^{-1}}$, with the maximum total during March 2013 and the minimum ~~estimate falling~~ total flux occurring in November 2012. The posterior totals ~~range~~ ranged between 149.5 and -375.1 $\mathrm{kt\ CO_2\ week^{-1}}$, with the maximum occurring in March 2013 and the minimum in October 2012. During the winter months, from March to July, the posterior ~~overlaps with the prior~~fluxes fell within the uncertainty limits of the prior fluxes. The posterior total ~~estimate moves~~ flux moved outside of the prior's confidence limits during August and September 2012,

which ~~is~~ was during the South African spring period. Posterior estimates ~~are larger compared to~~ were larger compared with the prior estimates. Data ~~are~~ were missing during October 2012 and from December 2012 ~~to~~ with January 2013, and therefore the estimates are completely overlapping during these periods. When observations ~~are~~ were available during the summer months in ~~October~~ November 2012 and February 2013 the posterior ~~estimates are larger compared to~~ fluxes were more positive compared with the prior estimates.

The total prior fossil fuel ~~estimate~~ flux was flat and close to 150 kt CO_2 $week^{-1}$ during the winter months, and close to 110 kt CO_2 $week^{-1}$ during the summer months. This stepped effect in the fossil fuel fluxes is due to the simple representation of the domestic emissions in the fossil fuel inventory. It is unlikely that fossil fuel emissions would have a sharp change between summer and winter~~, and therefore as~~. As a separate sensitivity analysis presented in a companion paper we adjusted the assumption of domestic emissions such that domestic emissions were distributed uniformly during the year. The inversion had the effect of reducing the total estimate, particularly during May 2012 and between March 2013 and June 2013, to a value of as low as 138 kt CO_2 $week^{-1}$ during the winter months and to a value of 107 kt CO_2 $week^{-1}$ during the summer months and early winter 2013. The ~~total estimates~~ posterior total fossil fluxes and confidence bounds for June 2013 ~~are~~ were outside of those for the prior estimates for the full month. Compared ~~to~~ with the total flux, the range of the total fossil fuel fluxes ~~is~~ was much narrower (between 100 and 160 kt CO_2 $week^{-1}$), and the confidence bounds around the estimates ~~are~~ were also narrower. This is not immediately apparent from the plot, but the range of the y-axis needed to be adjusted for the fossil fuel fluxes, otherwise it would appear as a thin line if plotted on the same range as the total fluxes.

~~Biogenic prior total fluxes range~~ Total prior NEE fluxes ranged between values close to zero and -494.9 kt CO_2 $week^{-1}$ and between zero and -483.1 kt CO_2 $week^{-1}$ for the posterior estimates. During the winter months posterior estimates were generally contained within the limits of the prior estimates, except for May 2012, where the total ~~estimate~~ NEE flux was slightly lower compared ~~to~~ with the prior. From August 2012 to September 2012, the posterior total ~~estimate was~~ NEE fluxes were well above the total prior ~~estimate~~estimates, indicating that the total uptake of CO_2 by the domain was reduced by the inversion during this period.

Comparisons of the ~~biogenic~~ NEE and fossil fuel ~~estimates~~ fluxes to the total estimates show that the variability in the total flux estimates was driven by variability in the ~~biogenic~~ NEE fluxes, and differences between the posterior and prior total estimates ~~are~~ were mainly driven by ~~changes to the biogenic~~ adjustments to the NEE fluxes induced by inversion. ~~The fossil fuel flux estimates are limited to a very narrow range compared to the total and biogenic fluxes, with much narrower confidence bands compared with biogenic fluxes. Since the inversion tended to increase biogenic fluxes when it was decreasing~~ As would be expected, due to the large uncertainty around the NEE prior fluxes, the inversion was not strongly constrained by these priors, and therefore mainly adjusted the NEE fluxes, rather than the fossil fuel fluxes, ~~it is evident that the inversion is most likely unable to correctly differentiate between these two fluxes when they co-occur in the same pixel. Therefore the inversion is correcting the prior biogenic fluxes to force the posterior modelled concentrations to better match with~~ so that the modelled concentrations better matched the observed concentrations~~, regardless of whether the correction is due to incorrectly specified biogenic or fossil fuel fluxes~~. This is unsurprising as we provided no information regarding what proportion of the observed CO_2 concentration was ~~from~~ attributable to fossil fuel contributions and what proportion was ~~from biogenic sources. The~~

posterior total flux estimate for a pixel and across the whole domain is more reliable compared to the individual fossil fuel and biogenic flux contributions. The cycle in the biogenic fluxestracks that of temperature well, with emissions more negative during the summer months. There is a lag between when the biogenic emissions begin to decrease and when temperatures begin to decrease, where temperatures start to drop in May but the biogenic emissions start to drop around Augustdue to NEE fluxes.

Temperatures are at their highest between January and March. This is due to the rainfall peaking in the winter months so that the periodof greatest productivity occurs in spring and early summer when temperatures are rising but still cool. The opposite trend throughout the rest of South Africa which is dominated by summer rainfall, where the driest period is usually around spring when temperatures start to rise quickly but water availability is at its lowest.

Prior (grey) and posterior (green) total weekly flux estimates () (right-side axis) and uncertainty limits (shaded area), represented as a 95 % confidence interval, across the full domain of Cape Town, where (a) is the total flux, (b) is the fossil fuel flux, and (c) is the biogenic flux. The daily temperature (°C) as recorded at the Cape Point GAW station is provided by the lower red line (right-side axis). The total weekly flux is obtained by summing over the working week, weekend and day and night-time sources from all pixels for each week.

3.3.1 Monthly Totals

Table ?? provides the total also the region's dry period. Both the prior and posterior flux estimates (NEE fluxes were most positive during this period, with NEE fluxes close to zero. The posterior NEE fluxes were more positive than the prior fluxes during these periods, i.e. the posterior estimates indicated less uptake of CO_2) for the domain over a one month period (four weeks) and their uncertainties expressed as standard deviations. The prior total flux estimates ranged between -1112.1 and 368.7 , whereas this range was narrowed for the posterior estimates to be between -896.8 and 433.9 . The estimates of the total fluxes demonstrate that the inversion has the ability to either decrease or increase the overall flux for the domain, allowing the domain to go from a prior net emitter of to a posterior net uptaker of and vice versa. For the months when the prior net flux for the domain was negative, the inversion had the impact of making the posterior net flux more positive . This occurred in August, September and November 2012 during the early summer months. The prior uncertainty in the total flux for the month ranged between 51.1 and 96.4 , with larger uncertainties during the summer months. The uncertainties of the posterior estimates were reduced by between 8.7% and 39.9% to range between 35.4 and 81.4 .

The inversion acted to reduce the total fossil fuel flux for all months. The prior total fossil fuel flux was approximately 597 during the winter months and 432 during the summer months. The posterior estimates were reduced by at most 24 , with the largest reductions in May to June 2012 and March to June 2013, which was during the winter periods. The prior and posterior uncertainty in the total estimates were very similar for all months, at approximately 2.8 for prior estimates and 2.7 for posterior estimates.The largest uncertainty reduction was 16.4% for June 2013. Both June 2012 and November 2012 had uncertainty reductions over 10%.

In general, the inversion had a dampening effect on the total biogenic flux, resulting in posterior estimates that were closer to zero. The prior total NEE ranged between -1544.4 and -229.6 , with the largest uptake predicted for early summer (September

and November 2012). The lowest uptake was predicted for March 2012 and 2013. The posterior estimates ranged between -1274.4 and -175.2 , with the lowest amount of uptake also taking place in March 2012 and 2013. In May 2012, June 2012 and April 2013 the inversion led to a more negative total biogenic flux, indicating larger uptake of predicted by the posterior estimates by biogenic processes compared with the prior estimates. The uncertainties for the biogenic fluxes were similar in magnitude compared with those of the total flux estimates, with uncertainty reductions of between 8.6 and 40.0%.

The total estimates were obtained for only those pixels which had an assigned prior fossil fuel flux. As the inversion was unlikely to distinguish between fossil fuel and biogenic fluxes from the same pixel, this total was obtained in order to assess how the innovations applied to the biogenic emissions by the inversion altered the total flux for the pixels contributing towards the fossil fuel emissions. The total flux from the fossil fuel pixels was made larger in August, September and November 2012 and February 2013, which is during the spring and summer seasons. Between March and July 2012 and between March and June 2013 the total estimates were made smaller by the inversion, which occur during the winter months. The largest decrease in the total flux occurred in May 2012, where the prior estimate was reduced from 387.4 to a posterior estimate of 198.5 minimum temperatures occurred between July and October, and during this period the NEE estimates were negative, indicating CO_2 . Similarly large reductions occurred between March 2013 and June 2013. The largest increase in total flux over the fossil fuel emitting pixels uptake. The peak in the uptake by biogenic processes occurred in September 2012, with an increase of 183.8 . Here the inversion changed a large negative prior net total flux over the fossil fuel pixels into a near-zero posterior net flux. to early November. This was during the spring period, after the winter rainfall period, and when the temperatures in the region begin to increase. The posterior estimates were less negative than the prior NEE fluxes, indicating that the NEE fluxes estimated by CABLE may have overestimated the amount of CO_2 uptake.

3.3.1 Model Assessment

To determine if the prescribed prior covariance parameters were consistent with the model assumptions, the sum of the squared normalised residuals was compared against the χ^2 distribution. For most months the standardised statistic was close to one, but in the case of June and July 2012, this statistic was above 2. We did not scale the variances independently for each month, and therefore the single scaling factor of 2 for the prior flux variances was not large enough for all months. The statistic remained below 2.5 for all months, and had a minimum of 0.68 for the month of November 2012. The mean of the statistic over all months was 1.48. A subsequent study will assess an alternative approach to determine prior flux uncertainties, which would guarantee compliance of the sum of the squared normalised residuals to follow the χ^2 distribution.

4 Discussion and Conclusions

4.1 Inversion Innovation

When comparing the differences between the observed and modelled concentrations for the reference inversion, these residuals are much smaller when using the posterior flux estimates compared with the prior estimates. Moreover, the The inversion was

able to substantially improve the agreement between the prior and posterior modelled concentrations, with posterior modelled concentrations ~~were able to track~~ tracking most of the local events observed in the measurements. The ~~agreement between the observed and modelled concentrations was also greatly improved, with standard deviations of the posterior residuals smaller by more the 1 compared with those of the observed concentrations. Corrections to the sources~~ most notable corrections to the pixel-level fluxes by the inversion were ~~mainly made to the pixel~~ made to those with the largest ~~point industrial source~~industrial point sources, to pixels located on Robben Island where ~~known, unaccounted for emission activities~~ activities unaccounted for in the inventory were taking place, and to the areas dominated by ~~biogenic sources, which had the greatest amount of uncertainty associated with the prior estimates relative to other sources included in the inversion~~NEE fluxes and located relatively close to the measurement sites. This evidence suggests that the inversion ~~setup~~ framework used here has had some success in capturing information regarding the CO_2 ~~emissions~~ fluxes in the CT domain, and has applied reasonable corrections to the sources considered.

The inversion was able to reduce uncertainty of the total flux within a pixel by up to 97.7%, and was able to reduce the uncertainty in the total ~~monthly~~ weekly flux over the whole domain by up to ~~39.9%. In particular, the uncertainty of the total flux attributed to pixels expected to contain fossil fuel sources was reduced to be between 3.0 and 11.6%.~~ 50.5%. The largest innovation to a fossil fuel ~~source~~ flux was applied to a pixel which contained an important point source in the domain ~~, which was~~ - a crude oil refinery. This facility can process up to 100,000 barrels of crude oil per day. Unlike most industrial sources in the area which would be expected to have fairly consistent activity from day to day, activity at the crude oil refinery would depend on oil supply and on the global oil prices. During the period of March 2012 to June 2013, the global monthly oil price deviated between $117.29 in March 2012 and $90.73 in June 2012, ending on $99.74 in June 2013 (World Bank commodity prices). In addition, the consumption of liquid petroleum gas and heavy furnace oils may have decreased during this period (City of Cape Town, 2015). As this is a source with a large amount of expected variability, it ~~'s~~ is unlikely that the inversion method with distant measurement sites will be able to adequately estimate the flux in this pixel. In order for this to occur, the measurement site would have to consistently view the source during periods of both high and low activity in order to provide an unbiased estimate~~of the mean flux from the site which is representative of the average activity occurring at the site. Inverse modelling~~ . An inversion could be used to estimate this particular source if a ring of instruments were placed around the site in order to capture information from the site at all times, regardless of the prevailing wind direction, such as suggested for the Otway CO_2CRC carbon capture project (Cook, 2012).

Compared ~~to~~ with the fossil fuel emissions, relative innovations to the ~~biogenic~~ NEE fluxes were much larger, due to the large uncertainty prescribed to these fluxes~~, particularly to the~~ . The largest innovations were made to natural areas near the CBD of CT, as well as to agricultural regions within the domain, particularly those close to the measurement sites. The prior estimates are dependent on the CABLE land-atmosphere interaction model ~~, anddespite being dynamically~~ and, although driven by the CCAM regional climate model, which ~~relies on~~ in turn was driven by reliable reanalysed observations of the climate from NCEP, is still under ongoing development for use over South Africa. There is a great deal of uncertainty in its ability to simulate fluxes over the fynbos biome, as there is for most dynamic vegetation models (Moncrieff et al., 2015). In general, the inversion tended to increase the ~~biogenic~~ NEE fluxes so that the total flux was less negative compared ~~to~~ with the priors, ~~except~~

~~in the case of May to July 2012, where the overall flux was made more negative. This occurs during the winter period, which is the dominant growing season in the Western Cape. Therefore the CABLE model appears to overestimate the drawdown of in the warmer, drier period, but underestimates the drawdown in the cooler, wetter season~~indicating that the amount of productivity estimated by CABLE may be overestimated.

4.2 Distinguishing Fossil Fuel and Biogenic Emissions

The spatial distribution and magnitude of the posterior fluxes and their uncertainties is strongly dependent on the prior spatial assignment of sources. In a city like CT, ~~these two sources are usually overlapping~~fossil fuel and NEE fluxes are usually co-occurring in the same pixel, with vegetation within the city representing a significant and large sink of CO_2~~, particularly during the growing season. If~~ . Under the current framework, if no fossil fuel ~~emission source is thought~~ source is prescribed to occur in a particular pixel, the inversion ~~under the current setup will~~ would only be able to adjust the ~~biogenic emissions if in reality~~ NEE flux in this pixel, as the fossil fuel flux and its uncertainty are set to zero. If there is an unknown fossil fuel source in ~~the pixel~~ a pixel the inversion would lead to a better match between the modelled and observed concentrations, but a worse NEE flux estimate. Therefore the success of the inversion is largely dependent on how well the spatial extent of fossil fuel and biogenic sources are prescribed in the prior information. ~~The fossil fuel and biogenic sources could be better differentiated if there were additional measurements of $\Delta^{14}C$ at each of the sites (Turnbull et al., 2015).~~

In Bréon et al. (2015) ~~biogenic fluxes were applied at~~ NEE fluxes were aggregated to a larger grid size ~~compared to~~ than the fossil fuel emissions. Effectively this means that perfect correlation was applied between the ~~biogenic~~ NEE fluxes for all pixels which fall within the same larger ~~grid cell~~NEE pixel. By distinguishing the biogenic and fossil fuel sources in this way, it may allow the inversion to correctly allocate corrections between the fossil fuel and biogenic sources. We attempted to implement a similar idea by allowing correlation between ~~biogenic emissions~~ NEE fluxes of neighbouring pixels and not prescribing correlations ~~with~~ between fossil fuel sources. As the model tended to reduce fossil fuel emissions and increase ~~biogenic~~ NEE fluxes in the same pixel, it appears that the inversion is ~~likely not correctly adjusting the individual sources~~unlikely to adequately adjust the individual fluxes making up the total flux from a pixel. With the large coverage of vegetation within the domain, it is unlikely that a measurement network with only two sites could accurately estimate a given industrial point source, but there is still potential to monitor the overall city emissions, and assess the feasibility of inventory information. This is evident from the large uncertainty reductions attained by the inversion for the total weekly flux estimates aggregated over the domain.

The posterior uncertainty of any linear combination of terms from the control vector of the fluxes (including the difference between two fluxes from the same pixel) will always be reduced or (at worst) left unchanged relative to the prior uncertainty of the same linear combination of elements (Jackson, 1979; Jackson and Matsu'ura, 1985). This means that although negative correlation between the flux components may be introduced through the inversion as observations are made of the sum of these components, the uncertainty in both the difference between fluxes from the same pixel and the total flux within a pixel will be reduced.

If we define the distinction between the fossil fuel flux and NEE flux within the same pixel i as the variance of the difference between the fossil fuel and NEE fluxes $s_{f,i} - s_{NEE,i}$, this will be equal to the sum of the variances of these two fluxes minus

twice the covariance between them: $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} - 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ will be negative. Therefore although the posterior uncertainty of the difference in these fluxes will always be larger than the sum of the individual posterior flux uncertainties, it will be smaller than the prior uncertainty of this linear combination of terms. The ability of the inversion under the current framework to distinguish between NEE and fossil fuel fluxes is limited as the posterior uncertainties are still large, and therefore the uncertainty of $s_{f,i} - s_{NEE,i}$ is large.

On the other hand, when we aggregate these fluxes from the same pixel to get the total flux within a pixel $s_{f,i} + s_{NEE,i}$, the uncertainty of this term is equal to $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} + 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ is negative. When we aggregate fluxes from the same pixel, the uncertainty of this total is smaller due to the both the smaller posterior uncertainties of the individual fluxes and also because the covariances are negative. This demonstrates that the value of the inversion is to reduce the uncertainty on each of the individual fluxes and to additionally reduce the uncertainty of the aggregation of the NEE and fossil fuel flux within the same pixel. In this case, the reduction in the uncertainty of the sum of fluxes within the same pixel is strongly dependent on the size of the uncertainty of the NEE flux, which is usually the larger uncertainty.

We have shown in the supplementary material (section 1.8) that if we reduce the uncertainty in the modelled concentrations, the negative off-diagonal covariance terms of $\mathbf{C_s}$ become larger in magnitude. To improve the ability of the inversion to estimate the total flux within a pixel, we need to improve the skill of the atmospheric transport model, and we need to reduce the uncertainty in the estimates of the NEE. As it stands, with a large prior uncertainty in the estimation of the NEE fluxes from the CABLE model which remains a large posterior estimate after the inversion, the distinction between the fossil fuel and NEE flux from the same pixel, $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} - 2 \times C_{s(f,i;NEE,i)}$, is not very different from the prior estimate, $C_{s_0(f,i;f,i)} + C_{s_0(NEE,i;NEE,i)}$.

4.3 Strengths and Limitations

This paper represents a first attempt at estimating CO_2 fluxes at the high resolution of 1 km by 1 km over CT, ~~distinguishing~~ solving for individual fossil fuel, terrestrial and oceanic biogenic sources. A previous network design study for South Africa (Nickless et al., 2015b), which aggregated the NEE and fossil fuel fluxes up to a 15 km×15 km resolution, showed the aggregation errors could be ~~very~~ high at the regional level. By avoiding this aggregation, and maintaining the 1 km by 1 km ~~through the~~ resolution of the atmospheric transport model throughout the inversion process, ~~aggregation errors could be minimised. We also used the CABLE model to provide estimates of the biogenic fluxes which was dynamically coupled to the CCAM regional climate model, and therefore prior emissions would be determined by the estimated climatic variables at this high resolution scale. This is advantageous over using modelled fluxes which are derived at a different scale to the atmospheric transport, and which then have to be scaled artificially to the resolution of the inversion~~we attempted to minimise these aggregation errors. Maintaining this resolution is computationally expensive, but possible due to the relatively small domain size of 100 km by 100 km.

A limitation of this study is that human respiration was not explicitly accounted for. With a population of over three million, this flux could represent up to ~~104.7~~26 kt CO_2 week^{-1}, if we attribute 1 kilogram CO_2 per day to each person (Bréon et al., 2015). This represents between a fifth to a quarter of the total fossil fuel flux estimated for the domain, and therefore is by no

means a negligible quantity. Including this fossil fuel flux in the inventory information would most likely lead to confounding between the domestic emissions and the human respiration, as ~~this source~~ these two sources would have been calculated based on population data. ~~A sensitivity analysis could be performed where the domestic emissions include an additional contribution from human respiration.~~ Domestic emissions were ~~also~~ heavily dependent on the assumption regarding how domestic heating emissions were distributed during the year. This had a large impact on the temporal profile of fossil fuel emissions, resulting in a lower average emission in summer compared ~~to~~ with winter, which persisted in the posterior estimates of the fossil fuel fluxes. ~~An additional sensitivity analysis in the companion paper considers assuming homogeneous heating emissions through the year, and investigates the impact on the total posterior estimates~~Consequently the prior uncertainties in the domestic emissions were larger compared with other fossil fuel sources. We have included sensitivity analyses on the specification of the fossil fuel emissions and their uncertainties in a companion paper.

This inversion was performed by solving for the ~~actual~~ total observed concentrations, rather than ~~using the gradient method of Bréon et al. (2015)~~solving for the gradient in concentration measurements between two sites (Bréon et al., 2015). There are several reasons why ~~this is~~ the gradient method would not have been suitable in these circumstances. Firstly, we had the advantage of a background site which viewed background levels over 70% of the time. As our city is located within the Southern Hemisphere, far less variability is expected between the hourly measurements or from week to week compared ~~to~~ with the Northern Hemisphere, and we would expect the ~~oceanic boundaries~~ CO_2 concentration at the oceanic boundaries of our domain to be similar to ~~have very similar~~ background levels of CO_2~~concentrations~~. This was confirmed by the results of the inversion, which made almost no adjustments to the oceanic boundary concentrations, but made slightly larger adjustments to the boundary concentrations ~~of~~ at the north and east ~~borders~~terrestrial boundaries. This implies that the adjustments made by the inversion were largely due to the surface flux sources within the domain~~, which then allowed for the stronger agreement between the posterior modelled and observed concentrations~~. In addition, there are no large expected sources located anywhere near the ~~boundary~~ boundaries of the domain. The next major city in the proximity of CT is Port Elizabeth, which is over 600 km away. For this reason, it is unnecessary to solve for the gradients between the two sites, as ~~in these circumstances we do~~ we did not require the removal of unknown outside sources from the observation dataset. This allowed us to use the entire measurement record, which is an advantage as we only had two CO_2 measurement sites available.

Secondly, the gradient method would likely have performed poorly here, as the direction of travel of an air parcel between the two sites would not necessarily be in a straight line due to the topography of the site and demonstrated by the plot of the ~~domain sensitivity~~ sensitivities at the two sites to the domain (Figure 16), as well as the modelled wind fields provided in supplementary material section 1.3. Therefore extracting observations based on the prevailing wind speed ~~at the site~~ and directions at the sites would have not represented true gradients in the CO_2 concentrations between the two measurements sites.

~~The sensitivity to the domain also reveals~~ These sensitivities of the sites to the surface fluxes also reveal that the sites are often viewing oceanic sources (Figure 16). A limitation of this study is that a single time series of ocean biogenic emissions was used as the prior estimate for all oceanic pixels in the domain. The fluxes from the near-coastal oceanic pixels are likely to have significant spatial heterogeneity, although smaller compared ~~to~~ with the terrestrial biogenic fluxes. A way of improving

this would be to use the output of a model representing atmosphere-land-ocean biogeochemical exchanges to provide prior fluxes over both the land and ocean. The CSIR's Variable Resolution Earth System Model (VRESM) is such a model currently under development, which aims to couple CCAM, CABLE, and CSIR's Variable-Cubic Ocean Model (VCOM), and Pelagic Interactions Scheme for Carbon and Ecosystem Studies (PISCES) to model ocean transport and biogeochemistry (Engelbrecht et al., 2016). Due to the amount of shipping activity around the CT harbour and within the ocean domain viewed by the Robben Island site, the inventory analysis could ~~also include~~ be improved by the inclusion of information on shipping routes so that the inversion can adjust fossil fuel ~~sources~~ fluxes in these ocean pixels as well.

The uncertainty in the ~~biogenic sources~~ NEE fluxes played an important role in the outcome of the inversion. If tighter uncertainty limits could be placed on the ~~biogenic~~ NEE flux estimates from the land-atmosphere exchange model, ~~made possible~~ it would allow the inversion to better distinguish between the fossil fuel and NEE fluxes. This could be attained by validation work through ~~, for example,~~ eddy-covariance flux measurements over dominant vegetation types within the domain, ~~it would allow the inversion to better constrain the emissions from the fossil fuel sources~~for example. The use of a land-atmosphere exchange model well suited to the vegetation within the domain of a ~~city scale~~ city-scale inversion is essential for ~~reducing the tendency~~ improving the ability of the inversion to adjust the ~~most uncertain sources to suit the observations, rather than~~ fossil fuel sources which are of foremost interest. The dependence on knowing the NEE fluxes well in order to estimate the fossil fuel fluxes could be reduced if there were additional measurements of Δ^{14}C and δ^{13}C isotope measurements at each of the sites, including at the background site (Turnbull et al., 2015; Newman et al., 2016).

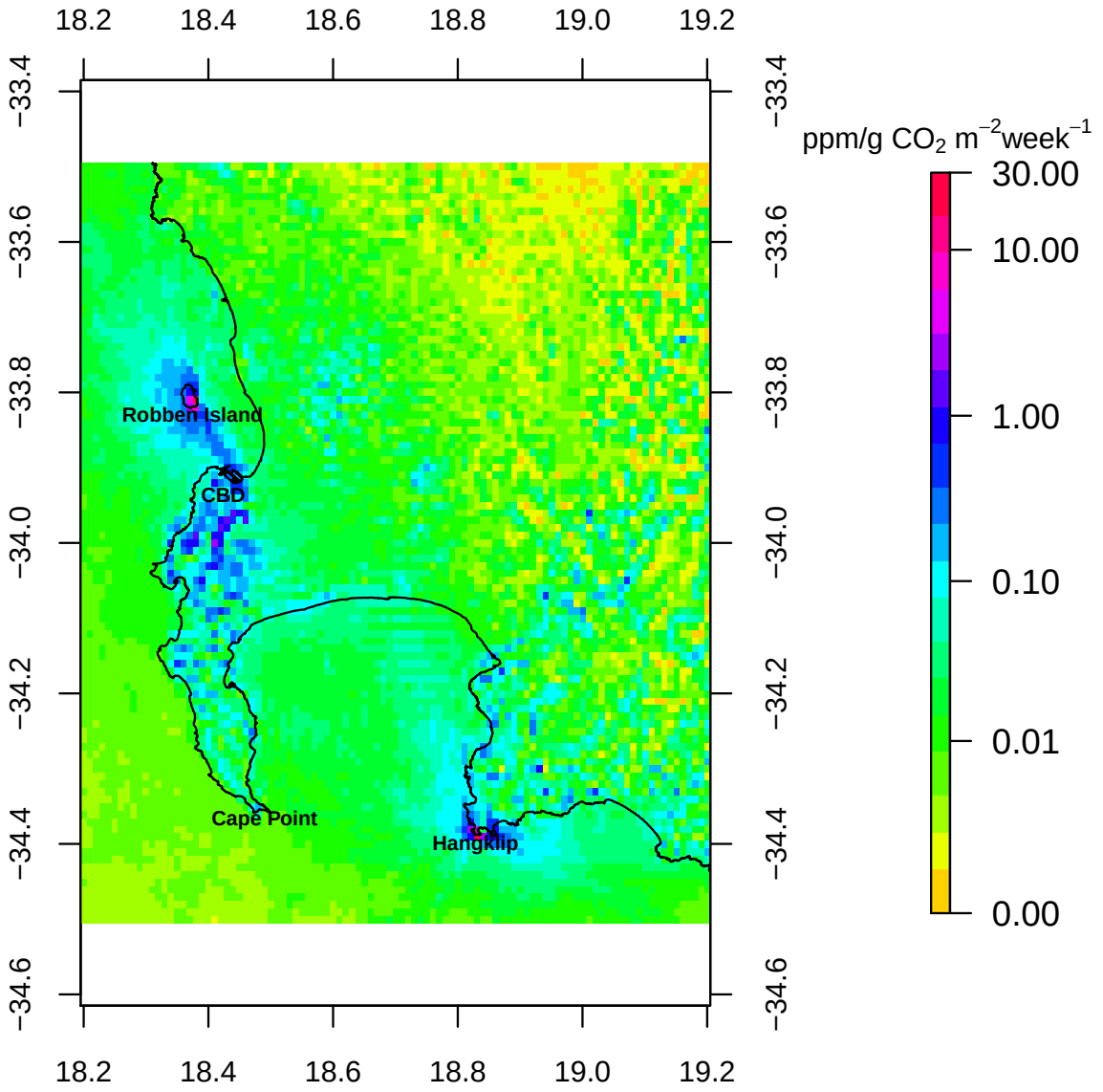


Figure 16. Mean ~~weakly~~ weekly sensitivities ($\mathrm{ppm\,kg^{-1}\ CO_2 m^{-2}\,week^{-1}}$) of the measurement sites at Robben Island and Hangklip to ~~the~~ each surface ~~sources~~source pixel, ~~which are derived from the mean sensitivities over all sensitivity matrices for the full duration of the study~~plotted on a log scale.

5 Conclusions

We have presented the results of a city-scale atmospheric inversion for Cape Town, South Africa. We have shown that the current inversion framework was able to reduce the misfit between the observed and modelled concentrations by making reasonable adjustments to the fossil fuel and NEE fluxes. We were able to reduce the total uncertainty in CO_2 weekly flux from the domain by up to 50%. We have demonstrated the ability of an inversion to obtain an improved estimate of an aggregated flux, even when the uncertainty in one component is large. To realistically use this framework for an operational inversion system for use in MRV, the uncertainty around the NEE estimates for Cape Town needs to be reduced. Further qualifying concentration observations according to the contributions from anthropogenic sources and those from the biosphere will help the inversion to disentangle the corrections to these two fluxes, and will reduce the dependency of the fossil fuel flux corrections on the uncertainty of the NEE flux estimates.

Author contributions. AN installed and maintained all the instrumentation at Robben Island and Hangklip, obtained the measurements and processed these into hourly concentrations, ran and processed the result of the LPDM in Fortran, produced all code and ran the inversion in Python, processed all the inversion results using R Statistical Software, produced all graphics and tables, and was responsible for the development of the manuscript which forms part of her PhD. PJR was the main scientific supervisor, oversaw all implementation of the inversion, and provided guidance on the presentation of results. E-GB provided the Cape Point concentration and meterological data, and provided assistance and guidance in the installation and maintenance of the Picarro instrumentation. FE provided the CCAM and CABLE data. BE provided guidance on statistical issues. RJS provided guidance on the location of the sites and provided advice on the interpretation of the biogenic fluxes. All authors had the opportunity to comment on the manuscript.

Competing interests. The authors declare that they have no conflict of interest

Code and data availability. Data and code related to the Bayesian inversion procedure can be made available upon request

Acknowledgements. We would like to acknowledge and thank Dr. Casper Labuschagne and Danie van der Spuy of the South African Weather Service for their assistance in maintaining the instruments at Robben Island and Hangklip, and Dr. Casper Labuschagne for his guidance on processing the instantaneous CO_2 concentration data; Martin Steinbacher for providing guidance and schematics on the calibration system used on the Picarro instruments; Robin Poggenpoel and Jacobus Smith of Transnet for allowing us access to the lighthouses; Peter Saaise of Transnet, the Robben Island lighthouse keeper, (and his daughter) for assisting when the instrument was not responding; Marek Uliasz for providing us access to the code for his LPDM model; Thomas Lauvaux for providing guidance on processing the LPDM results and useful discussion on the boundary contribution in the inversion; Keith Moir, Rooi Els, for providing wind data near Hangklip. Use was made of the University of Cape Town ICTS-HPC cluster. Please see http://hpc.uct.ac.za/ for details. We would like to thank Andrew Lewis of

the University of Cape Town HPC facility for providing useful advice on improving the efficiency of the Python runs. This research was funded by competitive parliamentary grant funding from the Council of Scientific and Industrial Research awarded to the Global Change Competency Area towards the development of the Variable-resolution Earth System Model (VRESM; Grants EEGC030 and EECM066). Additional funding was obtained from the South African National Research Foundation for the Picarro instrumentation.

References

Andres, R. J., Boden, T. A., Bréon, F. -M., Ciais, P., Davis, S., Erickson, D., Gregg, J. S., Jacobson, A., Marland, G., Miller, J., Oda, T., Olivier, J. G. J., Raupach, M. R., Rayner, P., and Treanton, K.: A synthesis of carbon dioxide emissions from fossil-fuel combustion, Biogeosciences, 9, 1845–1871, doi: 10.5194/bg-9-1845-2012, 2012.

Andres, R. J., Boden, T. A., and Higdon, D.: A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission, Tellus B, 66, 23616, doi: 10.3402/tellusb.v66.23616, 2014.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bourquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., Zhu, Z.: TransCom 3 inversion intercomparison: impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi: 10.1029/2004GB002439, 2006.

Bellassen, V. and Stephan, N.: Accounting for carbon: Monitoring, reporting and verifying emissions in the climate economy, Cambridge University Press, Cambridge, UK, 2015.

Boden, T. A., Marland, G., and Andres, R. J.: Global, regional, and national fossil fuel CO_2 emissions, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A., doi: 10.3334/CDIAC/00001_V2011, 2011.

Boon, A., Broquet, G., Clifford, D. J., Chevallier, F., Butterfield, D. M., Pison, I., Ramonet, M., Paris, J. -D., and Ciais, P.: Analysis of the potential of near-ground measurements of CO_2 and CH_4 in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model, Atmos. Chem. Phys., 16, 6735–6756, doi:10.5194/acp-16-6735-2016, 2016.

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual atmospheric CO_2 sources and sinks: 1. Method and control inversion, J. Geophys. Res., 104, 26161–26178, 1999.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707–1724, doi: 10.5194/acp-15-1707-2015, 2015.

Brioude, J., Angevine, W. M., Ahmadov, R., Kim, S. -W., Evan, S., McKeen, S. A., Hsie, E. -Y., Frost, G. J., Neuman, J. A., Pollack, I. B., Peischl, J., Ryerson, T. B., Holloway, J., Brown, S. S., Nowak, J. B., Roberts, J. M., Wofsy, S. C., Santoni, G. W., Oda, T., and Trainer, M.: Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NOxand CO_2 and their impacts, Atmos. Chem. Phys., 13, 3661–3677, doi: 10.5194/acp-13-3661-2013, 2013.

Brunke, E. -G, Labuschagne, C., Parker, B., Scheel, H. E, and Wittlestone, S.: Baseline air mass selection at Cape Point,South Africa: application of ^{222}Rn and other filter criteria to CO_2, Atmos. Environ., 38, 5693–5702, doi:10.1016/j.atmosenv.2004.04.024, 2004.

Carslaw, D. C. and Ropkins, K.: openair — an R package for air quality data analysis, Environ. Model. Softw., 27–28, 52–61, doi:10.1016/j.envsoft.2011.09.008, 2012.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morgui, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res., 115, D21307, doi: 10.1029/2010JD013887, 2010.

Chouinard, C., Béland, M., McFarlane, N.: A simple gravity wave drag parameterization for use in medium-range weather forecast models, Atmos. Ocean, 24, 91–110, 1986.

Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., and Ramonet, M.: Atmospheric inversions for estimating CO_2 fluxes: methods and perspectives, Climatic Change, 103, 69–92, 2010.

City of Cape Town: State of energy and energy futures report. Cape Town: City of Cape Town, http://www.capetown.gov.za/en/EnvironmentalResourceManagement/publications/Documents/State_of_Energy_+_Energy_Futures_Report_2011_revised_2012-01.pdf, 2011.

City of Cape Town: State of energy report 2015. Cape Town: City of Cape Town, https://africancityenergy.org/getfile.php?id=22&category=0, 2015.

P. Cook (Ed.): Geologically storing carbon: The CO2CRC Otway Project: Learning from the Otway Project experience, CSIRO Press, Melbourne, p. 384, 2014.

Crosson, E.:A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, Appl. Phys. B, 92, 403–408, doi:10.1007/s00340-008-3135-y, 2008.

Davis, K. J., Deng, A., Lauvaux, T., Miles, N. L., Richardson, S. J., Sarmiento, D. P., Gurney, K. R., Hardesty, R. M., Bonin, T. A., Brewer, W. A., Lamb, B. K., Shepson, P. B., Harvey, R. M., Cambaliza, M. O., Sweeney, C., Turnbull, J. C., Whetstone, J., and Karion, A.: The Indianapolis Flux Experiment (INFLUX): A test-bed for developing urban greenhouse gas emission measurements, Elem. Sci. Anth., 5:21, doi: 10.1525/elementa.188, 2017.

UK Department for Environment, Food and Rural Affairs (Defra): UK ship emissions inventory. Final report. London: Crown, http://uk-air.defra.gov.uk/assets/documents/reports/cat15/1012131459_21897_Final_Report_291110.pdf, 2010.

UK Department for Environment, Food and Rural Affairs (Defra): Government GHG conversion factors for company reporting: Methodology paper for emission factors, London: Crown, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224437/pb13988-emission-factor-methodology-130719.pdf, 2013.

UK Department for Environment, Food and Rural Affairs (Defra): Treatment of uncertainties for national estimates of greenhouse gas emissions, http://uk-air.defra.gov.uk/reports/empire/naei/ipcc/uncertainty/contents.html, 2013.

Dlugokencky, E. and Tans, P.: Trends in atmospheric carbon dioxide, National Oceanic and Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL), available at: http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html, last access: 29 September 2016, 2016.

Duren, R. M. and Miller, C. E.: Measuring the carbon emissions of megacities, Nat. Clim. Change, 2, 560–562, doi: 10.1038/nclimate1629, 2012.

Engelbrecht, F. A., McGregor, J. L. and Engelbrecht, C. J.: Dynamics of the conformal-cubic atmospheric model projected climate-change signal over southern Africa, Int. J. Climatol., 29, 1013-1033., doi: 10/1002/joc.1742.29., 2009.

Engelbrecht, F. A., Landman, W. A., Engelbrecht, C. J., Landman, S., Bopane, M. M., Roux, B., McGregor, J. L., and Thatcher, M.: Multi-scale climate modelling over Southern Africa using variable-resolution global model, Water Research Commission 40-Year Celebration Conference, Kempton Park, 31 August - 1 September 2011, doi: 10.4314/wsa.v37i5.2, 2011.

Engelbrecht, C. J., Engelbrecht, F. A. and Dyson, L. L.: High-resolution model-projected changes in mid-tropospheric closed-lows and extreme rainfall events over southern Africa. Int. J. Climatol., 33, 173–187, doi: 10/1002/joc.3420, ~~2013~~ 2013.

Engelbrecht, F., Adegoke, J., Bopape, M-J., Naidoo, M., Garland, R., Thatcher, M., McGregor, J., Katzfey, J., Werner, M., Ichoku, C. and Gatebe, C.: Projections of rapidly rising surface temperatures over Africa. Env. Res. Letters., 10(8), 085004, doi: 10.1088/1748-9326/10/8/085004, 2015.

Engelbrecht, F. A., McGregor, J. L., Thatcher, M., Garland, R., Sovara, M., Bopane, M. M., and van der Merwe, J.: The Variable-Resolution Earth System Model and its simulations of the Benguela upwelling system. The International Conference on Regional Climate CORDEX 2016, Stockholm, Sweden, 17–20 May 2016, 2016.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge Univ. Press, New York, 2002.

Erickson, P. and Tempest, K.: Advancing climate ambition: Cities as partners in global climate action, Produced by Stockholm Environment Institute (SEI) in support of the UN Secretary-General's Special Envoy for Cities and Climate Change and C40, 2014.

Exbrayat, J. -F., Pitman, A. J. Abramowitz, G. and Wang, Y, -P.: Sensitivity of net ecosystem exchange and heterotrophic respiration to parameterization uncertainty, J. Geophys. Res. Atmos., 118, 1640–1651, doi:10.1029/2012JD018122, 2013.

Fawcett, R. A., Pitcher, G., Bernard, S., Cembella, A., and Kudela, R.: Contrasting wind patterns and toxigenic phytoplankton in the southern Benguela upwelling system, Mar. Ecol. Prog. Ser., 348, 19–31, doi: 10.3354/meps07027, 2007.

Feng, S., Lauvaux T., Newman, S., Rao, P., Ahmadov, R. Deng, A., Díaz-Isaac, L. I., Duren, R. M., Fischer, M. L., Gerbig, C., Gurney, K. R., Huang, J., Jeong, S., Li, Z., Miller, C. E., O'Keeffe, D., Patarasuk, R., Ser, S. P., Song, Y., Wong, K. W., and Yung, Y. L.: Los Angeles megacity: a high-resolution land–atmosphere modelling system for urban CO_2 emissions, Atmos. Chem. Phys., 16, 9019–9045, doi: 10.5194/acp-16-9019-2016, 2016.

Ganesan, A. L., Rigby, M., Zammit-Mangion, A., Manning, A. J., Prinn, R. G., Fraser, P. J., Harth, C. M., Kim, K. -R., Krummel, P. B., Li, S., Mühle, J., O'Doherty, S. J., Park, S., Salameh, P. K., Steele, L. P., and Weiss, R. F.: Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, Atmos. Chem. Phys., 14, 3855-3864, doi:10.5194/acp-14-3855-2014, 2014.

Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rödenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe – Part 1: mapping the atmospheric CO_2 signals, Atmos. Chem. Phys., 7, 3461-3479, doi: 10.5194/acp-7-3461-2007, 2007.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S. and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, J. Geophys. Res., 108(D24), 4756, doi: 10.1029/2002JD003018, 2003.

Gregor, L. and Monteiro P. M. S.: Is the southern Benguela a significant regional sink of CO2? S. Afr. J. Sci., 109(5/6), Art. #0094, 5 pages, doi: 10.1590/sajs.2013/20120094, 2013.

Gurney. K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 405, 626–630, 2002.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Kowalczyk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.: TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, Tellus B, 55, 555–579, 2003.

Gurney, K. R., Mendoza, D. L., Zhou, Y. Y., Fischer, M. L., Miller, C. C., Geethakumar, S., Du Can, S. D.: High resolution fossil fuel combustion CO_2 emission fluxes for the United States. Environ. Sci. Technol., 43 (14), 5535–5541, doi: 10.1021/es900806c, 2009.

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., and Abdul–Massih, M.: Quantification of fossil fuel CO_2 emissions on the building/street scale for a large U.S. city, Environ. Sci. Technol., 46, 12194–12202, doi: 10.1021/es3011282, 2012.

Holtslag, A. A. M. and Boville, B. A.: Local versus non-local boundary layer diffusion in a global climate model, J. Clim., 6, 1825–1842, doi: 10.1175/1520-0442(1993)006<1825:LVNBLD>2.0.CO;2, 1993.

Hutyra, L., Duren, R., Gurney, K. R., Grimm, N., Kort, E. A., Larson, E., Shrestha, G.: Urbanization and the carbon cycle: Current capabilities and research outlook from the natural sciences perspective. Earth's Future, 2: 473–495. doi: 10.1002/2014EF000255, 2014.

Intergovernmental Panel on Climate Change (IPCC): Good practice guidance and uncertainty management in national greenhouse gas inventories. Montreal: IPCC, 93–102, http://www.ipcc-nggip.iges.or.jp/public/gp/english/, 2000.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K. and Meyer, L.A. (Eds.)]. IPCC, Geneva, Switzerland, 151 pp, 2014.

Jackson, D. D.: The use of a priori data to resolve non-uniqueness in linear inversion, Geophys. J. R. astr. Soc., 57, 137–157, 1979.

Jackson, D. D and Matsu'ura, M.: A Bayesian approach to nonlinear inversion, J. Geophys. Res. 90 (B1), 581–591, 1985.

Janssens-Maenhout, G., Pagliari, V., Guizzardi, D., and Muntean, M.: Global emission inventories in the Emission Database for Global Atmospheric Research (EDGAR) – Manual (I). Gridding: EDGAR emissions distribution on global gridmaps. Joint Research Centre, Luxembourg: European Union, 33 pp, doi: 10.2788/81454, 2012.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K., C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds R., Jenne, R., Joseph, D.: The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteorol. Soc., 77, 437–472, 1996.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three dimensional global inverse model of the atmospheric transport, 2. Inversion of the transport of CO_2 in the 1980s, J. Geophys. Res., 104, 18555–18581, 1999.

Kaminski, T., Rayner, P J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4705–4715, 2001.

Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO_2 emissions: minimum site location requirements for the Los Angeles megacity, J. Geophys. Res., 118, 1–8, doi: 10.1002/jgrd.50135, 2013.

Kowalczyk, E. A., Garratt, J. R. and Krummel, P. B.: Implementation of a soil-canopy scheme into the CSIRO GCM - regional aspects of the model response, CSIRO Div. Atmospheric Research, Melbourne, Australia, Tech Paper No. 32, 59 pp., 1994.

Kowalczyk, E. A., Wang, Y. P. and Law, R. M.: CSIRO Atmosphere Biosphere Land Exchange model for use in climate models and as an offline model, CSIRO Marine and Atmospheric Research technical paper xxV ISBN 1 921232 39 0, 2006.

Lacis, A. A. and Hansen, J. E.: A parameterization for the absorption of solar radiation in the earth's atmosphere, J. Atmos. Sci., 31, 118–133, doi: 10.1175/1520-0469(1974)031<0118:APFTAO>2.0.CO;2, 1974.

Landman, S., Engelbrecht, F. A, Dyson, L., Engelbrecht, C., J., and Landman, W., A.: A short-range ensemble prediction system for South Africa, Water SA, 38(5), 765–774, doi: 10.4314/wsa.v38i5.16, 2012.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. Phys., 8, 3459–3471, doi: 10.5194/acp-8-3459-2008, 2008.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, doi: 10.5194/bg-6-1089-2009, 2009.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337–354, doi: 10.5194/acp-12-337-2012, 2012.

Lauvaux, T., Miles, N. L., Richardson, S. J., Deng, A., Stauffer, D. R., Davis, K. J., Jacobson, G., Rella, C., Calonder, G. -P., and DeCola, P. L.: Urban emissions of CO_2 from Davos, Switzerland: the first real-time monitoring system using atmospheric inversion technique, J. Appl. Meteorol. Climatol., 52, 2654–2668, doi: 10.1175/JAMC-D-13-038.1, 2013.

Lauvaux, T., Miles, N. L., Deng, A. Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), J. Geophys. Res. Atmos., 121, 5213–5236, doi: 10.1002/2015JD024473, 2016.

Law, R. M., Chen, Y., Gurney, K. R., and Transcom 3 Modellers: TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, Tellus B, 55, 580–595, 2003.

Law, R. M., Raupach, M. R., Abramowitz, G., Dharssi, I., Haverd, V.,Pitman, A. J.,Renzullo, L., Van Dijk, A. and Wang, Y. -P.: The Community Atmosphere Biosphere Land Exchange (CABLE) model Roadmap for 2012-2017, CAWCR Technical Report No. 057, 2012.

Le Quéré, C., Andres, R. J., Boden, T. A., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D.,Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, Earth Syst. Sci. Data, 5, 165–185, doi: 10.5194/essd-5-165-2013 2013.

McGregor, J. L., Gordon, H. B., Watterson, I. G., Dix, M. R., and Rotstayn, L. D.: The CSIRO 9-level atmospheric general circulation model, CSIRO Div. Atmospheric Research Tech, Paper No. 26, 89 pp., 1993.

McGregor, J. L.: Semi-Lagrangian advection on conformal-cubic grids. Mon. Weather Rev., 124, 1311–1322, doi: 10.1175/1520-0493(1996)124<1311:SLAOCC>2.0.CO;2, 1996.

McGregor, J. L. and Dix, M. R.: The CSIRO conformal-cubic atmospheric GCM, in: IUTAM Symposium on Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics, Limerick, Ireland, 2–7 July 2000, edited by: Hodnett, P. F., Kluwer, Dordrecht, 197–202, 2001.

McGregor, J. L.: A new convection scheme using a simple closure, in: Current issues in the parameterization of convection, BMRC, Melbourne, Australia, Research Report 93, 33-36, 2003.

McGregor, J. L.: Geostrophic adjustment for reversibly staggered grids. Mon. Weather Rev., 133, 1119–1128, doi: 10.1175/MWR2908.1, 2005a.

McGregor, J. L.: C-CAM: Geometric aspects and dynamical formulation. CSIRO Atmospheric Research Technical Paper, No 70, 41, 2005b.

McGregor, J. L. and Dix, M. R.: An updated description of the Conformal-Cubic Atmospheric Model, in: High Resolution Numerical Modelling of the Atmosphere and Ocean, edited by: Hamilton, K. and Ohfuchi, W., Springer, New York, USA, 51–76, 2008.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, Proc. Natl. Acad. Sci. U. S. A., 109(22), 8423–8428, doi: 10.1073/pnas.1116645109, 2012.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 110, D24107, doi: 10.1029/2005JD005970, 2005.

Moncrieff, G.R., Scheiter,S. Slingsby, J. A. and Higgins, S. I.: Understanding global change impacts on South African biomes using Dynamic Vegetation Models, S. Afr. J. Bot., 101, 16–23, doi: 10.1016/j.sajb.2015.02.004, 2015.

Newman, S., Xu, X., Gurney, K. R., Hsu, Y. K., Li, K. F., Jiang, X., Keeling, R., Feng, S., O'Keefe, D., Patarasuk, R., Wong, K. W., Rao, P., Fischer, M. L., and Yung, Y. L.: Toward consistency between trends in bottom-up CO_2 emissions and top-down atmospheric measurements in the Los Angeles megacity, Atmos. Chem. Phys., 16, 3843–3863, doi: 10.5194/acp-16-3843-2016, 2016.

Nickless, A., Scholes, R. J. and Filby, E.: Spatial and temporal disaggregation of anthropogenic CO_2 emissions from the City of Cape Town, S. Afr. J. Sci., 111(11/12), Art. $\#2014-0387$, 8 pages, doi: 10.17159/sajs.2015/20140387, 2015.

Nickless, A., Ziehn, T., Rayner, P. J., Scholes, R .J., and Engelbrecht, F.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 2: Sensitivity analyses and South African test case, Atmos. Chem. Phys., 15, 2051–2069, doi: 10.5194/acp-15-2051-2015, 2015.

NRC (Committee on Methods for Estimating Greenhouse Gas Emissions): Verifying greenhouse gas emissions: methods to support international climate agreements (9780309152112), The National Academies Press, Washington, D.C., 2010.

Oda, T., Lauvaux, T., Lu, D., Rao, P., Miles, N. L., Richardson, S. J. and Gurney, K. R.: On the impact of granularity of space-based urban CO_2 emissions in urban atmospheric inversions: A case study for Indianapolis, IN, Elem Sci Anth, 5, 28, doi: 10.1525/elementa.146, 2017.

Peylin, P., Baker, D., Sarmiento, J., Ciais, P., and Bousquet, P.: Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO_2 data, J. Geophys. Res., 107, 4385, doi: 10.1029/2001JD000857, 2002.

Rayner, P. J., Enting, I. G., Francey, R. J., and Langenfelds, R. L.: Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, Tellus B, 51, 213–232, 1999.

Rayner, P. J., Law, R. M, Allison, C. E., Francey, R. J., Trudinger, C. M., Pickett-Heaps, C.: Interannual variability of the global carbon cycle (1992-2005) inferred by inversion of atmospheric CO_2 and $\delta^{13}C$ measurements, Global Biogeochem. Cy., 22, GB3008, doi: 10.1029/2007GB003068, 2008.

Rotstayn, L. D.: A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I: Description and evaluation of the microphysical processes, Q. J. R. Meteorol. Soc., 123, 1227–1282, 1997.

Roux, B.: Ultra high-resolution climate simulations over the Stellenbosch wine producing region using a variable-resolution model, MSc Thesis, Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa, 106 pp., 2009.

Schmidt, F.: Variable fine mesh in spectral global model, Beitr. Phys. Atmos., 50, 211–217, 1977.

Schuh, A. E., Lauvaux, T., West, T. O., Denning, A. S., Davis, K., J., Miles, N., Richardson, S., Uliasz, M., Lokupitiya, E., Cooley, D., Andrews, A., and Ogle, S.: Evaluating atmospheric CO_2 inversions at multiple scales over a highly inventoried agricultural landscape, Glob. Change Biol., 19, 1424–1439, doi: 10.1111/gcb.12141, 2013.

Schwarzkopf, M. D. and Fels, S. B.: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes, J. Geophys. Res., 96, 9075–9096, 1991.

Shiga, Y. P., Michalak, A. M., Gourdji, S. M., Mueller, K. L., and Yadav, V.: Detecting fossil fuel emissions patterns from subcontinental regions using North American in situ CO_2 measurements, Geophys. Res. Lett., 41(12), 4381–4388, doi: 10.1002/2014GL059684, 2014.

Shrout, P. E. and Fleiss, J. L.: Intraclass correlations: uses in assessing rater reliability, Psychol. Bull., 2, 420–428, 1979.

Seibert, P. and Frank, A.: Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode, Atmos. Chem. Phys., 4, 51–63, doi: 10.5194/acp-4-51-2004, 2004.

South African Department of Energy: Digest of South African energy statistics, Pretoria: Department of Energy. http://www.energy.gov.za/files/media/explained/2009%20Digest%20PDF%20version.pdf, 2009.

Statistics South Africa: Census 2011 statistical release, P0301.4., Pretoria: Statistics South Africa, 2011.

Staufer, J., Broquet, G., Bréon, F. -M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO2 emissions based on atmospheric inversion, Atmos. Chem. Phys., 16, 14703–14726, doi: 10.5194/acp-16-14703-2016, 2016.

Strong, C., Stwertka, C., Bowling, D. R., Stephens, B. B. and Ehleringer, J. R.: Urban carbon dioxide cycles within the Salt Lake Valley: A multiple-box model validated by observations, J. Geophys. Res., 116, D15307, doi: 10.1029/2011JD015693, 2011.

Seto, K. C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J. E., Müller, D. B., Murakami, J., Nagendra, H., and Ramaswami, A.: Human settlements, infrastructure and spatial planning, in: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlömer, S., von Stechow, C., Zwickel, T., and Minx, J. C., Cambridge, United Kingdom and New York, NY, USA, 2014.

Sugar, L. and Kennedy, C.: A low carbon infrastructure plan for Toronto, Canada, Can. J. Civ. Eng., 40, 86–96, doi: 10.1139/cjce-2011-0523, 2013.

Tans, P. and Keeling, R.: Mauna Loa CO_2 monthly mean data, NOAA/ESRL www.esrl.noaa.gov/gmd/ccgg/trends/ and Scripps Institution of Oceanography scrippsco2.ucsd.edu/, 2016.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, 2005.

Thatcher, M. and McGregor, J. L.: Using a scale-selective filter for dynamical downscaling with the conformal cubic atmospheric model, Mon. Weather Rev., 137, 1742–1752, 2009.

Thatcher, M. and McGregor, J. L.: A technique for dynamically downscaling daily-averaged GCM datasets over Australia using the Conformal Cubic Atmospheric Model, Mon. Weather Rev., 139, 79–95, 2010.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., and Patarasuk, P.: Toward quantification and source sector identification of fossil fuel CO_2 emissions from an urban area: Results from the INFLUX experiment, J. Geophys. Res. Atmos., 120, 292–312, doi: 10.1002/2014JD022555, 2015.

Uliasz, M.: The atmospheric mesoscale dispersion modeling system, J. Appl. Meteorol., 31, 139–149, 1993.

Uliasz, M.: Lagrangian particle modeling in mesoscale applications, in: Environmental Modelling II, Computational Mechanics Publications, Southampton, UK, 71–102, 1994.

UN–Habitat: Cities and climate change: global report on human settlements 2011, Earthscan, United Nations Human Settlements Programme (UN–Habitat), 2011.

Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., van Gorsel, E. and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, J. Geophys. Res., 116, G01034, doi: 10.1029/2010JG001385, 2011.

Whittlestone, S., Kowalczyk, E., Brunke, E. G., and Labuschagne, C.: Source Regions for CO_2 at Cape Point Assessed by Modelling 222Rn and Meteorological Data, Technical Report for the South African Weather Service, Pretoria, South Africa, 2009.

Wu, L., Bocquet, M., Chevallier, F., Lauvaux, T., and Davis, K.: Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions, Tellus B, 65, 20894, doi: 10.3402/tellusb.v65i0.20894, 2013.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO_2 emissions?, Atmos. Chem. Phys., 16, 7743–7771, doi: 10.5194/acp-16-7743-2016, 2016.

Zhang, L., Zhang, H. and Li, Y.: Surface energy, water and carbon cycle in China simulated by the Australian community land surface model (CABLE), Theor. Appl. Climatol.,96(3),375–394, doi:10.1007/s00704-008-0047-z, 2009.

Zhang, H ., Pak, .B., Wang, Y. P., Zhou, X., Zhang, Y. and Zhang, L.: Evaluating surface ~~qater~~ water cycles simulated by the Australian community land surface model (CABLE) across different spatial and temporal domains. J. Hydrometeor, 14, 1119–1138, 2013.

Ziehn, T., Nickless, A., Rayner, P. J., Law, R. M., Roff, G., and Fraser, P.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 1: Methodology and Australian test case, Atmos. Phys. Chem., 14, 9363–9378, doi: 10.5194/acp-14-9363-2014, 2014.

Estimates of CO_2 fluxes over the City of Cape Town, South Africa, through Bayesian inverse modelling

Alecia Nickless[1,2], Peter J. Rayner[3], Francois Engelbrecht[4], Ernst-Günther Brunke[5], Birgit Erni[1], and Robert J. Scholes[6]

[1]Department of Statistical Sciences, University of Cape Town, Cape Town, 7701, South Africa
[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, OX2 6GG, UK
[3]School of Earth Sciences, University of Melbourne, Melbourne, VIC 3010, Australia
[4]Climate Studies and Modelling and Environmental Health, CSIR, Pretoria, 0005, South Africa
[5]South African Weather Service c/o CSIR, P.O. Box 320, Stellenbosch, 7599, South Africa
[6]Global Change Institute, University of the Witwatersrand, Johannesburg, 2050, South Africa

Correspondence to: Alecia Nickless alecia.nickless@phc.ox.ac.uk

Abstract. We present a city-scale inversion over Cape Town, South Africa. Measurement sites for atmospheric CO_2 concentrations were installed at Robben Island and Hangklip lighthouses, located downwind and upwind of the metropolis. Prior estimates of the fossil fuel fluxes were obtained from a bespoke inventory analysis where emissions were spatially and temporally disaggregated and uncertainty estimates determined by means of error propagation techniques. Net ecosystem exchange (NEE) fluxes from biogenic processes were obtained from the land atmosphere exchange model CABLE (Community Atmosphere Biosphere Land Exchange). Uncertainty estimates were based on the estimates of net primary productivity. CABLE was dynamically coupled to the regional climate model CCAM (Conformal Cubic Atmospheric Model), which provided the climate inputs required to drive the Lagrangian particle dispersion model. The Bayesian inversion framework included a control vector where fossil fuel and NEE fluxes were solved for separately.

Due to the large prior uncertainty prescribed to the NEE fluxes, the current inversion framework was unable to adequately distinguish between the fossil fuel and NEE fluxes, but the inversion was able to obtain improved estimates of the total fluxes within pixels and across the domain. The median of the uncertainty reductions of the total weekly flux estimates for the inversion domain of Cape Town was 28%, but reach as high as 50%. At the pixel level, uncertainty reductions of the total weekly flux reached up to 98%, but these large uncertainty reductions were for NEE-dominated pixels. Improved corrections to the fossil fuel fluxes would be possible if the uncertainty around the prior NEE fluxes could be reduced. In order for this inversion framework to be operationalised for monitoring, reporting and verification (MRV) of emissions from Cape Town, the NEE component of the CO_2 budget needs to be better understood. Additional measurements of $\Delta^{14}C$ and $\delta^{13}C$ isotope measurements would be a beneficial component of an atmospheric monitoring programme aimed at MRV of CO_2 for any city which has significant biogenic influence, allowing improved separation of contributions from NEE and fossil fuel fluxes to the observed CO_2 concentration.

1 Introduction

Cities are under pressure to reduce their carbon dioxide emissions. In the last 10 years (2006 to 2015), the mean annual increase in carbon dioxide (CO_2) concentrations in the global atmosphere has been 2.11 ppm per year (Dlugokencky and Tans, 2016) (NOAA/ESRL 2016), a sharper rise in CO_2 emissions than the preceding decades (IPCC, 2014). Approximately 76% of current anthropogenic greenhouse gas emissions are comprised of CO_2 contributions (IPCC 2014). While cities cover a mere 2% of the global land surface area, they are responsible for 70% of anthropogenic greenhouse gas emissions (UN–Habitat, 2011), and between 71 and 76% of CO_2 emissions from global final energy use (Seto et al., 2014). Annual urban CO_2 emissions are more than double the net terrestrial or ocean carbon sinks (Le Quéré et al., 2013).

South Africa is the single largest emitter of CO_2 on the continent of Africa, and the 13th largest emitter in the world (Boden et al., 2011). South African cities are home to 63% of the present population (Statistics South Africa, 2011), and by 2030 this is predicted to be 71%. The population of Cape Town (CT) has been rising at 2.5% per annum over the past two decades, and currently is nearly 4 million (City of Cape Town, 2011). Cities are seen as having the greatest potential to provide solutions for emissions reduction and climate change mitigation (Seto et al., 2014; Wu et al., 2016). By reducing the CO_2 impact of cities, cities play a pivotal role in decreasing their own climate vulnerability. But there are also additional co-benefits which include improving air-quality, energy access, public health, city liveability, and developing the economy and job creation through advances in green technology (Seto et al., 2014).

Formal climate action plans are developed by governments and city managers whereby the roadmap for implementing greener policies is provided, such as encouraging and developing public transport which makes use of low emission technologies, mass and rapid transport systems, and building retrofits (Sugar and Kennedy, 2013; Erickson and Tempest, 2014). Many cities are taking it on themselves to respond to the climate crisis, reacting to limited international and national policy progress (Hutyra et al., 2014). But to determine if the plans implemented are having the anticipated effect of lowering CO_2 emissions, monitoring is required. Monitoring, reporting and verification (MRV) is a concept which is fundamental to most market and policy-based mechanisms in climate economics (Bellassen and Stephan, 2015). In order for emission reduction strategies to be properly implemented and assessed, an MRV approach should be adopted so that emission reduction claims can be validated in a consistent and reliable manner. Currently, the primary source of this information for cities is by means of emissions inventories. This relies on the collection of activity data to provide an inventory of emissions from different sectors or specific point sources. These inventories are not perfect representations of CO_2 emissions. They are heavily dependent on accurate reporting, emission factors, and on assumptions regarding temporal or spatial disaggregation of emissions (Andres et al., 2012), where errors associated with these emission estimates increase with higher spatial and temporal resolutions (Andres et al., 2014). As the importance of these inventories increases due to the need to quantify emissions and assess emission targets, it has become necessary to verify the accuracy of these estimates (NRC, 2010). Adequate MRV implementation requires transparency, quality and comparability of information, with narrow uncertainty estimates (Wu et al., 2016). Currently, uncertainties associated with urban emissions far exceed emission reduction goals, and therefore verification remains challenging. The large amount of uncertainty is due to factors such as incomplete data, inconsistency in reporting between different institutions or facilities,

fugitive emissions from point sources such as those caused by gas leaks, and methodology which is rarely checked against scientific standards and procedures (Hutyra et al., 2014). A way of verifying inventory data for a city, and reducing uncertainty of inventory estimates, is by means of the Bayesian atmospheric inversion technique. This method aims to take advantage of continuous measurement of CO_2 concentrations from a network of atmospheric monitoring sites located in and around a city. By attempting to model the CO_2 concentrations at these sites, the inversion is able to provide corrections to the inventory of CO_2 emissions from the city, so that the mismatch between the modelled and observed concentrations is reduced.

Several regional or mesoscale atmospheric inversions have been published (Lauvaux et al., 2008, 2009, 2012; Schuh et al., 2013), and more recently city-scale inversion studies have been conducted in Europe and North America (Strong et al., 2011; Duren and Miller, 2012; McKain et al., 2012; Brioude et al., 2013; Kort et al., 2013; Lauvaux et al., 2013; Bréon et al., 2015; Turnbull et al., 2015; Boon et al., 2016; Oda et al. , 2017). These top down approaches make use of an atmospheric transport model to relate observations of CO_2 concentrations in the atmosphere to the CO_2 fluxes from the domain of interest (Lauvaux et al., 2012). This method applies corrections to the inventory data, which enters the inversion calculation by means of the prior estimates. This paper reports the results for an atmospheric inversion for CT, South Africa.

Making use of point measurements of CO_2 concentrations means that the effects of all fluxes of CO_2 are observed as an aggregated total. It is challenging to separate out these aggregated CO_2 fluxes into different components of the total CO_2 budget without additional measurements, such as Δ^{14}C (Turnbull et al., 2015) and δ^{13}C isotope measurements (Newman et al., 2016), or without high confidence in the spatial and temporal patterns of fluxes (Shiga et al., 2014). Even when additional measurements of CO_2 mole fractions are available, at the current point in time, background atmospheric conditions are not sufficiently characterised to use isotope tracers to discriminate between fossil fuel and biogenic fluxes (Turnbull et al., 2015). To conduct a Bayesian atmospheric inversion at the city-scale, a detailed CO_2 inventory analysis is required, where all the main contributors to the anthropogenic CO_2 budget are considered. Apart from their use in an atmospheric inversion, better understanding of the underlying processes at the urban scale and improved quantification of CO_2 emissions provides information contributing towards the policy decisions made by urban practitioners, helps to improve understanding of urban dynamics, and informs future scenarios (Hutyra et al., 2014). An example of this is the detailed street level inventory analysis undertaken in the Hestia project for U.S. cities Indianapolis, Los Angeles, Phoenix and Salt Lake City (Gurney et al., 2012; Davis et al., 2017). Preceding these inventories was the Vulcan inventory which covers the contiguous U.S. (Gurney et al., 2009). These detailed inventories have made possible atmospheric inversion exercises, as well as other top down methods for obtaining urban CO_2 flux estimates, for these cities (Strong et al., 2011; Brioude et al., 2013; Bréon et al., 2015; Lauvaux et al., 2016). Such a detailed inventory analysis is not available for any South African city, and therefore a detailed spatially and temporally disaggregated inventory analysis of direct CO_2 emissions was undertaken for CT specifically for the use of this atmospheric inversion exercise (Nickless et al., 2015a).

Atmospheric inversions have various sources of uncertainty, which include atmospheric transport modelling errors (particularly at night when the planetary boundary layer is shallow) (Geels et al., 2007); incorrect characterisation of prior flux estimates and their uncertainties (which includes errors in the inventory analysis) (Bréon et al., 2015; Lauvaux et al., 2016); atmospheric measurement errors (Gerbig et al., 2003); representation errors due to the comparison of a concentration mea-

surement at a point with a modelled concentration representative of a surface grid box (Gerbig et al., 2003); and aggregation errors which occur as fluxes from various sources are coerced into homogeneous grid cells (Kaminski et al., 2001). In the case of cities, atmospheric transport modelling is further complicated by small-scale turbulence, highly heterogeneous surface characteristics, and urban heat island effects (Hutyra et al., 2014; Bréon et al., 2015).

Therefore, careful consideration of the atmospheric transport model (or models) is required for an atmospheric inversion. The atmospheric transport modelling in this study was provided by the Conformal Cubic Atmospheric Model (CCAM) (McGregor and Dix, 2008) at the resolution of $1\,\mathrm{km}\times1\,\mathrm{km}$. CCAM, at a slightly coarser resolution, has already been used for a regional network design study over South Africa, making use of a similar Bayesian inversion framework (Nickless et al., 2015b), and has been verified over South Africa and over the CT target region at a spatial resolution of up to $1\,\mathrm{km}\times1\,\mathrm{km}$ (Roux, 2009; Engelbrecht et al., 2009, 2011).

High resolution inversions are required to quantify emissions down to the sector or point source level. Lauvaux et al. (2016) performed an ultra high resolution inversion where sector specific anthropogenic emissions were considered, but ignored biogenic fluxes. This was possible due to the selection of the dormant period for the inversion, when fluxes due the biosphere would have been at a minimum. When considering longer periods, or for cities in regions which may not have a dormant vegetation period, this assumption will not be valid, particularly for a medium-sized city, where natural processes can be a significant contributor to the carbon budget. Such would be the case for South African cities, such as CT and Johannesburg, where large national parks and other natural areas are located near or within city limits and within city vegetation growth is non-negligible. CT is also surrounded by a large agricultural sector consisting of vineyards and fruit orchards. Ironically, there are features of cities which allow for better plant growth. For example, the urban heat island effect leads to a longer growing season for plants, and reduced wind within cities leads to less plant stress resulting in better plant growth (Buyantuyev and Wu 2012). In addition, nitrogen deposition within cities leads to increased nutrient availability, and particularly in arid regions, cities cause augmented water availability for plants (Hutyra et al., 2014). If allowed growing space, plants can make a significant contribution to the carbon budget of a city.

Therefore, for a city like CT, biogenic fluxes cannot be ignored, and within atmospheric inversion studies are usually estimated by means of a land surface exchange model (Bréon et al., 2015; Staufer et al., 2016). Bréon et al. (2015) and Staufer et al. (2016) made use of the C-TESSEL land atmosphere scheme which is used in the ECMWF forecasting system. In this study we have made use of the CABLE (Community Atmosphere Biosphere Land Exchange) model to represent the biogenic $\mathrm{CO_2}$ fluxes in the $\mathrm{CO_2}$ budget (Kowalczyk et al., 2006). CABLE had the same spatial and temporal resolution as the meteorology. The average weekly fluxes for each pixel were calculated and used as the prior biogenic fluxes.

We present a Bayesian inversion framework used to obtain estimates of $\mathrm{CO_2}$ fluxes over CT, and present the results of the reference atmospheric inversion for a sixteen month period from March 2012 until June 2013. The domain considered was a $100\,\mathrm{km}\times100\,\mathrm{km}$ region with CT at the centre. The spatial resolution of the atmospheric transport model was set at $1\,\mathrm{km}\times1\,\mathrm{km}$, and the spatial resolution of the surface fluxes was made to match this resolution. Fluxes were solved for at a weekly time step, separately for day and night. Fossil fuel and biogenic fluxes were solved for separately, and fossil fuel fluxes separated into week and weekend fluxes.

2 Methods

2.1 Bayesian Inverse Modelling Approach

The Bayesian synthesis inversion method, as described by Tarantola (2005) and Enting (2002), was used to solve for the fluxes in this study. This method has been described for global inversions (Bousquet et al., 1999; Kaminski et al., 1999; Rayner et al., 1999; Gurney et al., 2002; Peylin et al., 2002; Gurney et al., 2003; Law et al., 2003; Baker et al., 2006; Rayner et al., 2008; Ciais et al., 2010), as well as for many of the recent city-scale inversions (Lauvaux et al., 2016; Bréon et al., 2015). The observed concentration (c) at a measurement station at a given time can be expressed as the sum of different contributions from the surface fluxes, from the domain boundaries, and from the initial concentration at the site. Concentrations at the measurement site can be modelled as:

$$\mathbf{c}_{mod} = \mathbf{H}\mathbf{s} \tag{1}$$

where $\mathbf{c}_{mod}$ are the modelled concentrations and $\mathbf{s}$ are various sources, where sources are any part of the domain which can provide a positive or negative contribution of CO_2. $\mathbf{H}$ is the Jacobian matrix representing the first derivative of the modelled concentration at the observational site and dated with respect to the coefficients of the source components (Enting, 2002). It provides the sensitivity of each observation to each of the unknown sources, where the sources can be either fluxes or concentrations of CO_2. Estimates of the unknown sources can be obtained by minimising the following cost-function with respect to $\mathbf{s}$:

$$J(\mathbf{s}) = \frac{1}{2}\left((\mathbf{c}_{mod}-\mathbf{c})^T {\mathbf{C}_\mathbf{c}}^{-1}(\mathbf{c}_{mod}-\mathbf{c}) + (\mathbf{s}-\mathbf{s_0})^T {\mathbf{C}_{\mathbf{s_0}}}^{-1}(\mathbf{s}-\mathbf{s_0})\right) \tag{2}$$

where $\mathbf{s}$ is the control vector of unknown surface fluxes and boundary concentrations we wish to solve for, $\mathbf{s_0}$ is the vector of prior flux and boundary concentration estimates, $\mathbf{C_c}$ is the uncertainty covariance matrix of the observations, and $\mathbf{C_{s_0}}$ is the uncertainty covariance matrix of the fluxes and boundary concentrations (Tarantola, 2005).

The solution to this minimisation problem is:

$$\mathbf{s} = \mathbf{s_0} + \mathbf{C_{s_0}}\mathbf{H}^T\left(\mathbf{H}\mathbf{C_{s_0}}\mathbf{H}^T + \mathbf{C_c}\right)^{-1}(\mathbf{c}-\mathbf{H}\mathbf{s_0}) \tag{3}$$

and the posterior covariance matrix can be determined as follows (Tarantola, 2005):

$$\mathbf{C_s} = \left(\mathbf{H}^T {\mathbf{C_c}}^{-1} \mathbf{H} + {\mathbf{C_{s_0}}}^{-1}\right)^{-1} \quad (4)$$

$$= \mathbf{C_{s_0}} - \mathbf{C_{s_0}} \mathbf{H}^T \left(\mathbf{H} \mathbf{C_{s_0}} \mathbf{H}^T + \mathbf{C_c}\right)^{-1} \mathbf{H} \mathbf{C_{s_0}}. \quad (5)$$

2.2 Control vector - $\mathbf{s}$

The control vector, $\mathbf{s}$, can be broken up into different components. The total CO_2 flux from a single surface pixel for a given week is made up of the following individual fluxes:

$$\mathbf{s}_{sf;\,i} = \mathbf{s}_{ff\ week\ day;\,i} + \mathbf{s}_{ff\ week\ night;\,i} + \mathbf{s}_{ff\ weekend\ day;\,i} + \mathbf{s}_{ff\ weekend\ night;\,i} + \mathbf{s}_{NEE\ day;\,i} + \mathbf{s}_{NEE\ night;\,i} \quad (6)$$

where $\mathbf{s}_{sf;\,i}$ is the total weekly surface flux from the i^{th} pixel, $\mathbf{s}_{ff\ week\ day;\,i}$ is the total fossil fuel flux during the day during the working week, $\mathbf{s}_{ff\ week\ night;\,i}$ is the total night-time fossil fuel flux during the working week, $\mathbf{s}_{ff\ weekend\ day;\,i}$ is the total weekend daytime fossil fuel flux, $\mathbf{s}_{ff\ weekend\ night;\,i}$ is the total weekend night-time fossil fuel flux, and $\mathbf{s}_{NEE\ day;\,i}$ and $\mathbf{s}_{NEE\ night;\,i}$ are the total day and night-time biogenic fluxes for the full week from the i^{th} pixel. The inversion solves for each of these separate fluxes. There are 101×101 = 10,201 surface pixels. Over the 16 month period from March 2012 to June 2013, separate monthly inversions are carried out for all months with sufficient valid concentration observations; a total of 13 inversions. Each monthly inversion solves for four weekly fluxes. Therefore a monthly inversion solves for 10,201×6×4 = 244,824 surface fluxes.

The mean day and night-time concentrations at each of the four domain boundaries for each week are the last components of the control vector. The inversion solves for 4×2×4 = 32 boundary concentrations (4 boundaries, day/night, 4 weeks). We solved for weekly concentrations at the boundaries as we expected these concentrations to show small changes on synoptic time scales, particularly inflow from the ocean boundaries. We avoided solving for too short a period so that the percentile filtering technique (see section 2.8) would never discard all measurements for a period. The maximum standard deviation in the hourly background CO_2 concentrations for a week was $0.8\,\mathrm{ppm}$.

As a sensitivity analysis, presented in a companion paper, we examined two alternative compositions of the control vector. We considered solving for a mean weekly flux for each month. In this case for a surface pixel we solved for two biogenic mean weekly fluxes (day and night) and four fossil fuel mean weekly fluxes (day and night working week, day and night weekend). We also considered a separate inversion for each week. In this case only the concentration measurements for one week were used and the individual weekly fluxes (two biogenic and four fossil fuel) were solved for, and this was repeated for each of the four weeks in the month. The benefit of these two alternative control vectors is that the resulting dimensions of the $\mathbf{C_{s_0}}$ matrix is much smaller compared with the reference case we present in this paper.

2.3 Concentration measurements - c

Two CO_2 monitoring sites were established at Robben Island and Hangklip lighthouses. Due to the dominant wind directions in CT (Fawcett et al., 2007), either from the south or north west, the location of the Robben Island and Hangklip stations were well suited for observing contributions from the area of interest, particularly from CT. The Hangklip site observed mainly background air, but occasionally viewed the biogenic-influenced continental air. Robben Island often observed air with enhancements from CT. The location of these sites in relation to the domain are shown in Figure 1. The average wind speed and direction across the domain, as modelled by CCAM, are shown in the supplementary material (supplementary material section 1.3).

Figure 1. Google Earth image of the domain, where Cape Town is located at the centre. The corner coordinates of the full domain are $33°29'42.00''$ south $18°11'42.00''$ east (top left), $33°29'42.00''$ south $19°12'18.00''$ east (top right), $34°30'18.00''$ south $18°11'42.00''$ east (bottom left), $34°30'18.00''$ south $19°12'18.00''$ east (bottom left). The locations of the measurement sites and the Cape Point GAW station background site are indicated together with images of these sites (Photo credits: Ernst Brunke and Alecia Nickless). CBD = central business district.

Each site was equipped with a Picarro Cavity Ring-down Spectroscopy (CRDS) (Picarro G2301) instrument. This instrument measures CO_2, methane (CH_4), and water vapour (H_2O) simultaneously, every five seconds, producing a precision of better than 0.05 parts-per-million volume (ppmv) for CO_2, 0.07 parts-per-billion volume (ppbv) for CH_4, and 100 ppmv for H_2O. This instrument maintains high linearity, precision, and accuracy over changing environmental conditions, requiring only minimal calibration, and is recognised as one of the highest precision instruments for measurement of the top three greenhouse gases (Crosson, 2008).

The inlet of the measurement tube at each site was located at the top of the lighthouse, and had a Gelman filter to prevent contamination of the instrument through aerosols or water droplets. The inlet tube led to a VICI rotary valve which directed the sampled air stream to the Picarro instrument. Approximately every four days the rotary valve switched to a calibration line which allowed the flow of calibration gas through the instrument for a period of half an hour.

The Robben Island lighthouse is an $18\,\mathrm{m}$ tall circular masonry tower, and the height of the focal plane of the light is $47\,\mathrm{m}$ above the high water level. The location of the lighthouse is $33°48'52.20''$ south and $18°22'29.25''$ east. The Hangklip lighthouse is a $22\,\mathrm{m}$ tall concrete tower, where the focal plane of the light is $34\,\mathrm{m}$ above the high water level. It is located at $34°23'11.40''$ south and $18°49'42.30''$ east. It is located on the tip of False Bay, opposite to Cape Point.

2.4 System Meteorology

CCAM is the variable-resolution global atmospheric model developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (McGregor, 1996; McGregor and Dix, 2001; McGregor, 2005a, b; McGregor and Dix, 2008). It employs a semi-implicit semi-Lagrangian method to solve the hydrostatic primitive equations. The Geophysical Fluid Dynamics Laboratory (GFDL) parameterisations for long-wave and short-wave radiation are used (Lacis and Hansen, 1974; Schwarzkopf and Fels, 1991), with interactive cloud distributions determined by the liquid and ice-water scheme of Rotstayn (1997). Total-variation-diminishing vertical advection is applied to solve for the advective process in the vertical. A stability-dependent boundary layer scheme based on Monin Obukhov similarity theory is employed (McGregor, 1993), together with the non-local treatment of the boundary layer scheme as described in Holtslag and Boville (1993). A canopy scheme is included, as described by Kowalczyk et al. (1994), having 6 layers for soil temperatures and soil moisture (solving Richard's equation) and 3 layers for snow. The cumulus convection scheme uses a mass-flux closure (McGregor, 2003), and includes downdrafts, entrainment and detrainment. Gravity wave drag is parameterised following Chouinard et al. (1986).

CCAM may be applied in stretched-grid mode to function as a regional climate model, thereby providing a flexible framework for downscaling reanalysis data or global circulation model simulations to high resolution over an area of interest. Stretched grids are obtained using the Schmidt (1977) transformation. A multiple-nudging approach was followed to downscale the $250\,\mathrm{km}$ resolution National Centres for Environmental Prediction (NCEP) reanalysis data (Kalnay et al., 1996) to a resolution of $60\,\mathrm{km}$ over southern Africa, $8\,\mathrm{km}$ over the south western Cape and subsequently to a $1\,\mathrm{km}$ resolution over the study area. The $8\,\mathrm{km}$ resolution domain stretched over an area of about $1300\times1300\,\mathrm{km}^2$, whilst the $1\,\mathrm{km}$ resolution domain centred over False Bay stretched over an area of about $160\times160\,\mathrm{km}^2$. Output was stored at a time resolution of 1 hour. CCAM was spectrally nudged with the synoptic-scale forcing reanalysis data at 6-hourly intervals for the period 1979-2013 using a

scale-selective Gaussian filter (Thatcher and McGregor, 2009, 2010). This forcing was applied from 900 hPa higher up into the atmosphere. Sea-surface temperatures from the NCEP data set were used as lower boundary forcing.

To justify the use of CCAM to provide modelled winds and other climatological variables, we rely on previous studies which have used this model for atmospheric transport modelling in our target area (Whittlestone et al., 2009), and studies which have validated CCAM at various spatial resolutions (Engelbrecht et al., 2009; Roux, 2009; Engelbrecht et al., 2011, 2013, 2015). In particular, CCAM has been able to satisfactorily recreate present-day rainfall totals and the rainfall seasonal cycle, as well as circulation patterns over South Africa (Engelbrecht et al., 2009), and has been able to simulate with some success mid-tropospheric closed-lows and extreme rainfall events (Engelbrecht et al., 2015). CCAM has been validated over the Stellenbosch wine-producing area, which falls within the domain of this inversion, with respect to temperature, relative humidity and wind speed at six different stations within this region (Roux, 2009). Those stations located within the high-resolution focus area of the stretched-grid obtained root mean square errors of 0.64 ms^{-1} or lower and correlations close to 1 between the modelled and observed wind speeds. Validating the wind product from CCAM further in a rigorous manner is beyond the scope of this paper.

2.5 Jacobian matrix - H

In order to generate the Jacobian matrix, $\mathbf{H}$, for the inversion procedure, which maps the surface fluxes and boundary inflows to the concentrations observed at the receptor sites, a Lagrangian particle dispersion model (LPDM) was run in backward mode. An LPDM simulates the release of a large number of particles from arbitrary receptors and records the location of the particles at fixed time steps (Uliasz, 1993, 1994). The model implemented in this study was developed by Marek Uliasz (1993), which will be referred to as LPDM. LPDM is driven by the hourly three-dimensional fields of mean winds (u, v, w), potential temperature and turbulent kinetic energy (TKE), which were obtained from the CCAM model. When LPDM is run backward in time, in receptor-orientated mode, the particle counts can be used to generate $\mathbf{H}$ for a given receptor site, as described in Ziehn et al. (2014) and Nickless et al. (2015b) following Seibert and Frank (2004).

The Jacobian for a 4 week period during each month of the study was generated by allowing the LPDM model to run in backward mode over a full 2 month period. Particle counts were extracted for the 4 weeks of interest. Particles were released every 20 seconds and each particle's position was recorded at 1 minute intervals. Particles that were near the surface were allocated to a surface grid box, corresponding to the surface pixels of the atmospheric transport model, and the total particle count within each of these boxes was determined. These counts depended on the dimensions and position of the surface grid boxes. The particle counts were used to calculate the source–receptor (s–r) relationship. We followed Seibert and Frank (2004) to convert the particle counts into the elements of the Jacobian matrix. As described in Ziehn et al. (2014), we modified the approach of Seibert and Frank (2004) to account for the particle counts which were produced by our LPDM as opposed to the mass concentrations which were output by the atmospheric transport model in their study. The resulting s–r relationship between the measurement site and source i at time interval n, which provide the elements of the matrix $\mathbf{H}$, is:

$$\frac{\partial \bar{\mathbf{c}_{\mathbf{sf}}}}{\partial \mathbf{s}_{in}} = \frac{\Delta T \mathrm{g}}{\Delta P} \overline{\left(\frac{N_{in}}{N_{tot}}\right)} \frac{44}{12} \times 10^3, \tag{7}$$

where $\bar{\mathbf{c}_{\mathbf{sf}}}$ is a volume mixing ratio (receptor) expressed in ppm and $\mathbf{s}_{in}$ is a mass flux density (source), N_{in} the number of particles in the receptor surface grid from source pixel i released at time interval n, ΔT is the length of the time interval, ΔP is the pressure difference in the surface layer, g is the acceleration due to gravity, and N_{tot} the total number of particles released during a given time interval.

In this inversion setup weekly fluxes of CO_2 were separated into day and night-time contributions, into fossil fuel and NEE contributions, and in the case of fossil fuels, into working week and weekend contributions. Therefore, to obtain the NEE contributions the particle count N_{in} was the sum over one week (ΔT=1 week (day/night)). For fossil fuel fluxes, the particle count was separated into the contribution from the working week and from the weekend, separately for day and night.

The surface layer height was set to $50\,\mathrm{m}$ which corresponds to approximately $595\,\mathrm{Pa}$ (ΔP). If we assume well mixed conditions, then the s–r relationship should be independent of the thickness of the surface layer, as long as the layer is not too deep, as the particle count will be adjusted proportional to the volume of the grid box. Under stable conditions, this may not be the case (Seibert and Frank, 2004). The spatial resolution of the surface flux grid boxes was set to be the same as that of the high-resolution subregion of the atmospheric transport model, resulting in a gridded domain consisting of 101×101 grid boxes (a resolution of approximately $0.01^\circ \times 0.01^\circ$ or $1\,\mathrm{km} \times 1\,\mathrm{km}$).

The fluxes from the surface pixels are expressed in $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ and are transformed through $\mathbf{H}$ into contributions to the concentration at the measurement site in units of ppm. The inversion solves for the concentrations at the boundary of the domain. Ziehn et al. (2014) shows that the Jacobian which provides the sensitivities of observed concentrations to boundary concentrations can be calculated as:

$$\frac{\partial \bar{\mathbf{s}_{\mathbf{B}}}}{\partial \mathbf{c}_b} = \frac{N_B}{N_{tot}} \tag{8}$$

where $\bar{\mathbf{s}_{\mathbf{B}}}$ is the concentration at the domain boundary, $\mathbf{c}_b$ is the volume mixing ratio, N_B is the number of particles from the specific domain boundary and N_{tot} the total number of particles viewed at the receptor site from any of the domain boundaries. The contribution to the observed concentration at the receptor site can be written as:

$$\mathbf{c}_b = \mathbf{H}_B \mathbf{s}_B \tag{9}$$

where $\mathbf{H}_B$ is the Jacobian with respect to the domain boundary concentrations, $\mathbf{s}_B$ the domain boundary concentrations and $\mathbf{c}_b$ the contributions from the boundary to the observed concentration at the measurement site in units of ppm. The row elements of $\mathbf{H}_B$ sum to one. Therefore the elements of $\mathbf{c}_b$ represent a weighted average of the concentrations at the domain boundaries,

and provide a basis concentration to which the contributions from the surface fluxes are added. Each inversion solves for 4 weekly domain boundary concentrations for each cardinal direction, separated by day and night.

2.6 Inventory of anthropogenic emissions

An inventory analysis was conducted specifically for this atmospheric inversion exercise (Nickless et al., 2015a). The anthropogenic emissions were subdivided into those due to road transport, airport and harbour emissions, residential lighting and heating, and industrial point sources. Road transport emissions were derived from modelled values of vehicle kilometres for each section of the road network, modelled from observed vehicle count data. The vehicle kilometres were scaled for each hour of the day, and reported separately for working week days and weekend days. Therefore the vehicle emissions for day and night are distinctive for the week / weekend and day / night periods.

Airport emissions were derived from landing and takeoff cycles, as reported by Airports Company South Africa for each month. We used the IPCC reported average emission factors for domestic and international fleets (IPCC, 2000), and these were used to convert the airport activity data into emissions of CO_2. Emissions were expected to be concentrated between 6:00 and 22:00, and so the monthly emission was divided evenly between these hours. Harbour emissions were derived for port activity published by the South African Ports Authority for each month. Based on the gross tonnage of vessels which docked at the port during the month, emissions could be derived as described in DEFRA (2010). The monthly emissions were divided equally between all hours of the month, as it was assumed that harbour activities would be continuous.

Residential emissions for lighting and heating were derived from population count data obtained for each of the municipal wards in 2011 (Statistics South Africa, 2011). The population of CT was 3,740,025, as reported in the 2011 census (Statistics South Africa, 2011). The South African government reports on the fuel used for domestic heating and lighting (South African Department of Energy, 2009). This was divided between the total population, and then allocated to each ward depending on the population residing in that area. The fuel usage was scaled according to the proportion of fuel used for cooking, lighting and heating, where 75% of the annual heating fuel usage was assumed to take place during the winter months (March to August). It was assumed that 75% of the annual energy consumed was used for heating, 20% for cooking and 5% for lighting.

CT provided monthly fuel usage by the largest industrial emitters. The reported fuel usage for the top fuel users were converted directly into CO_2 emissions by multiplying these figures with the Defra greenhouse gas emission factors (DEFRA, 2013a). The fuel types that were considered included heavy fuel oil, coal, diesel, paraffin and fuel gas which were divided into liquid petroleum gas and refinery fuel gas. As no information was available about when the activity was occurring at these facilities, the emissions were divided equally between all hours of the month.

Based on this inventory analysis, the percentage contribution of industrial point sources to the total fossil fuel emission for CT was 12.0%, 34.6% from vehicle road transport, 51.0% from the residential sector, and 2.4% from the airport and harbour transport. Residential emissions are a large contributor to the fossil fuel emission budget as well as one of the largest contributors to the uncertainties in the fossil fuel flux. This is due to the dependency that many people living in CT have on raw fossil fuel burning for heating and lighting. Emissions from power stations are a small component of the total fossil fuel flux from CT as the bulk of the direct emissions from power stations occur elsewhere in the country.

The total fossil fuel emissions for the domain were comparable with those from the EDGAR (Emission Database for Global Atmospheric Research) (v4.2) database (Nickless et al., 2015a). EDGAR is a global product on a $0.1^\circ \times 0.1^\circ$ grid, which provides the total anthropogenic emissions of CO_2 as estimated from proxy data such as population counts and information on the road transport network (Janssens-Maenhout et al., 2012). The total emissions estimated from our bespoke inventory analysis for 2012 were 22% higher relative to the emissions from EDGAR in 2010, where the emissions in our inventory tended to be concentrated over specific sources, such as oil refinery plants, whereas the EDGAR emissions were smoothed over the city region.

2.7 Biogenic emissions

CCAM was dynamically coupled to the land surface model CABLE, which allows for feedbacks between land surface and climate processes, such as leaf area feedback on maximal canopy conductance and latent heat fluxes (Zhang et al., 2013). This type of coupling has successfully been implemented in CSIRO's national earth system modelling scheme (Australian Community Climate and Earth System Simulator or ACCESS) and describes land-atmosphere exchanges of energy, carbon, and water using biogeochemical, vegetation-dynamic and disturbance processes (Law et al., 2012). Several studies have validated CABLE under different ecosystems and parameters using both global model simulations (e.g. Zhang et al. (2009); Wang et al. (2011)), and site level offline CABLE simulations (Exbrayat et al., 2013; Zhang et al., 2013).

The model produces hourly estimates of net ecosystem exchange (NEE), which were aggregated into weekly (day and night) flux estimates in units of $\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$, and used as the prior estimate of biogenic fluxes over the land surface. The spatial resolution of these prior NEE fluxes were kept at a $0.01^\circ \times 0.01^\circ$ resolution.

In terms of natural vegetation, the target domain is dominated by the fynbos biome. This biome is biodiverse, with many endemic species, and covers a relatively small area in South Africa, but a significant area within the domain of the inversion. The fynbos biome is poorly represented by dynamic vegetation models (Moncrieff et al., 2015). The land atmosphere exchange model CABLE was selected to couple with CCAM due to its use and development in regions of Australia which share similar characteristics to the savanna biome in South Africa, which has a coverage of over 50%. Its ability to simulate respiration and photosynthesis in the fynbos region is largely untested. In addition to the natural vegetation, a large agricultural sector is within the proximity of CT, consisting predominantly of vineyards and fruit orchards. The CT region experiences a Mediterranean climate with winter rainfall. Consequently, summers are hot and dry and winters are mild and wet. Therefore significant NEE fluxes take place during both winter and summer periods. The NEE in this region is limited by the amount of water availability, whereas temperatures are usually sufficiently high enough not to limit plant production and respiration.

The CO_2 fluxes over the ocean were obtained from Gregor and Monteiro (2013). This study characterised the seasonal cycle of air-sea fluxes of CO_2 in the southern Benguela upwelling system off the South African west coast. A time series of pCO_2, derived from total alkalinity and dissolved inorganic carbon and scatterometer-based wind, was obtained from six monthly cross-shelf cruises in the St. Helena Bay region during 2010. Daily CO_2 fluxes were derived from these pCO_2. These fluxes were applied as prior estimates to the ocean surface grids within the domain. Therefore, an assumption was made that ocean

CO_2 fluxes are relatively homogeneous in space near the south western coast of South Africa, but the inversion was given the ability to differentially adjust each of the ocean sources in the posterior estimates.

2.8 Domain boundary concentrations

The existence of the Cape Point Global Atmospheric Watch (GAW) station made CT an ideal candidate for a city-scale inversion exercise. The Cape Point station is located approximately $60\,\mathrm{km}$ south of CT within a nature reserve, situated on the southern-most tip of the Cape Peninsula at a latitude of $34°21'12.0''$ south and longitude of $18°29'25.2''$ east. The inlet is located on top of the $30\,\mathrm{m}$ measurement tower, which is located on a cliff $230\,\mathrm{m}$ above sea level. The station observes background measurements of CO_2 when observing maritime air advected directly from the south-western Atlantic Ocean. This is an extensive region stretching from $20°$ (sub-equatorial) to $80°$ (Antarctic region) (Brunke et al., 2004). Therefore, maritime measurements at Cape Point from the Southern Ocean are well representative of the background CO_2 signal influencing the Cape Peninsula. The background signal at Cape Point, obtained from a percentile filtering technique (Brunke et al., 2004), was used as the prior estimate of the concentrations at each of the four domain boundaries. The percentile filtering technique removes data influenced by the continent or anthropogenic emissions. Two 11-day moving percentiles, which are adjustable by tuneable factors, control the upper and lower threshold limits. This results in a subset of background measurements from Cape Point represented by a narrow concentration band contained within these limits. This filter, when applied to the Cape Point CO_2 measurements, selects approximately 75% of the data. The percentile-filtering technique has been shown to compare well with the more robust method of using contemporaneous radon (^{222}Rn) measurements to differentiate between marine and continental air.

This site provides a long term record of background CO_2 concentrations for the area. These continuous measurements of the background CO_2 levels meant that we were not dependent on the atmospheric transport model to produce estimates of CO_2 concentrations at the domain boundary, which are prone to large errors (Lauvaux et al., 2016). Due to the prevailing wind directions across the domain the "gradient approach" for solving for CO_2 was not appropriate. This gradient approach relies on the observed wind direction and wind speed to obtain a subset of the concentration measurements when the air flow is from one measurement site directly to another. The differential in the concentrations is modelled by the inversion (Lauvaux et al., 2013; Bréon et al., 2015; Staufer et al., 2016). Plots provided in the supplementary material (supplementary material section 1.3) show the average wind speed and direction for the domain for each month. In general, the wind direction was not favourable to the gradient approach, and with only two measurement sites, would have left little information to constrain the surface fluxes. When the wind is blowing from the south easterly direction, air from the Hangklip site curves northwards towards the interior and away from CT. When the Robben Island site is observing marine air on its way into the CT area from the Atlantic side, such as June 2013, the wind changes for the north westerly direction once it passes over CT to a more northerly direction, missing the Hangklip site.

The mean weekly background concentrations, separate for day and night, were determined from the percentile filtered measurements at the site, and were used as the prior domain boundary concentrations for each of the four cardinal directions. The inversion was then allowed to make small adjustments to these concentrations. The prior variance assigned to the boundary con-

centrations was equal to the variance of the measured hourly concentrations for that period. As the variability in the background CO_2 in the southern hemisphere is small, much smaller than for the northern hemisphere, this resulted in a tight constraint on the prior background CO_2 concentrations. Large adjustments by the inversion to the far-field domain boundary concentrations were not expected. The daytime weekly background concentrations are shown in Figure 2. The standard deviation in the hourly background CO_2 concentrations ranged between 0.32 and 0.90 ppm, with a mean of 0.62 ppm.

The boundaries of the domain were deliberately set to be far from the measurement sites so that contributions to the CO_2 concentration at a measurement site were dominated by the surface fluxes within the domain, rather than by the domain boundary concentrations.

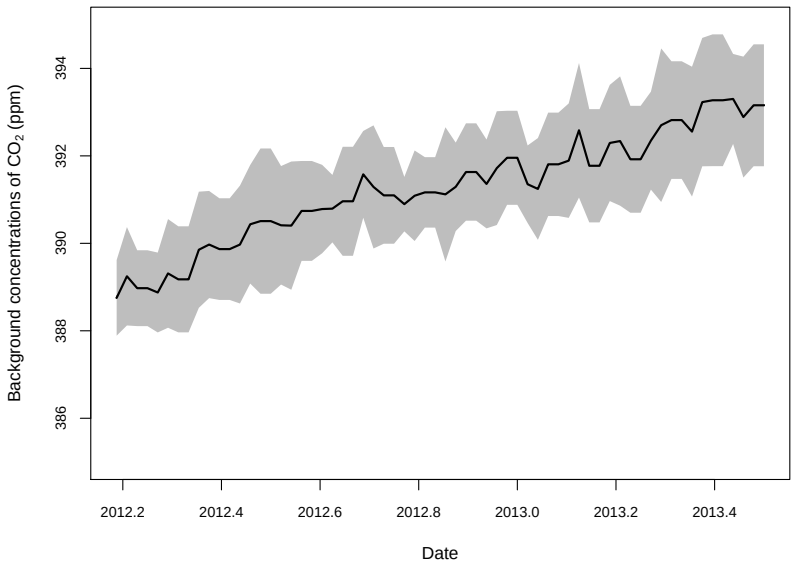


Figure 2. Weekly mean background concentrations of CO_2 (ppm) as measured at Cape Point GAW station, with 95% confidence interval represented by the grey shaded area. The mean concentrations are calculated from percentile filtered observations, extracting only those observations considered to be representative of background conditions.

2.9 Prior covariance matrix - $\mathbf{C_{s_0}}$

The uncertainty covariance matrix, $\mathbf{C_{s_0}}$, of the prior fluxes and domain boundary concentrations $\mathbf{s_0}$ determines in part how much freedom the inversion has to adjust these fluxes based on the observed concentrations $\mathbf{c}$. If the off-diagonal prior covariance elements are significantly different from zero, then the estimate for a each flux will be more dependant on the prior estimates of the surrounding fluxes compared with an inversion where the covariances between the uncertainties in the prior fluxes were set to zero. On the other hand, if the prior variances are large, the inversion is able to make large adjustments to flux estimates to obtain better agreement between the observed and modelled concentrations. The next two subsections explain how the original estimates of the uncertainties in the fluxes and observation errors were determined. The uncertainties in the prior fluxes were scaled by an additional factor of 2 to ensure goodness-of-fit of the covariance structure (see supplementary material section 1.2).

2.9.1 Fossil fuel emissions

Error propagation techniques were used to estimate the uncertainties in the sector specific fossil fuel emissions. This was described in Nickless et al. (2015a). An industrial point source flux $s_{0;ff}$ was derived from the equation

$$s_{0;ff} = AE \tag{10}$$

where A is the activity data, usually fuel usage, and E is the process-specific emission factor. The uncertainty in the flux was estimated from

$$C_{s_{0;ff}} = |s_{0;ff}|^2 \times ((\frac{\delta A}{A})^2 + (\frac{\delta E}{E})^2) \tag{11}$$

where $C_{s_{0;ff}}$ is the uncertainty in the flux estimate expressed as a variance, δA is the uncertainty in the activity data and δE the uncertainty in the emission factor, expressed as standard deviations. DEFRA (2013b) provides estimates of uncertainty in the activity data and emission factors under various industrial processes for each fuel type.

For vehicle emissions, which relied on count data, Poisson errors were assumed, and propagated together with the uncertainty in the conversion factors for the different vehicle types. For airport and harbour emissions, vessel counts were assumed to be correct, and therefore the uncertainty in the emissions contained within the emission factors for the different vessel types and activities. For aircraft, these errors are assumed to be 34% for the international fleet and 28% for the domestic fleet (IPCC, 2000). The error estimate for berth and manoeuvring activities of shipping vessels is reported to be between 20% and 30%, and therefore a conservative estimate of 30% was used (DEFRA, 2010). For domestic heating and lighting, the estimates relied on population census, which had a reported omission rate of 15%. There was no information available on the variability in fuel

usage between households, and therefore the uncertainty in the domestic emissions was set at 30% as a relatively arbitrary, but conservative level. Domestic emissions due to fossil fuel burning was a large contributor to the overall fossil fuel flux of the domain. As the percentage uncertainty assigned to these fluxes was large, uncertainties in the domestic emissions was a significant contributor to the overall uncertainty in the fossil fuel fluxes.

After accounting for the scaling of the uncertainty estimates to improve goodness-of-fit of the covariance structure, the resulting uncertainty estimates (expressed as standard deviations) ranged between 6.7% to 71.7% of the prior fossil fuel emission estimate, with a median percentage of 34.9% to 38.4% depending on the month. These values are in general more conservative compared with uncertainties that were determined by Bréon et al. (2015) for the AirParif inventory, which were set at 20% throughout. The spatial distribution of the fossil fuel fluxes during the month of March 2012 are mapped in Figure 3. The daytime fossil fuel emissions have a mean of $0.006\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$ and go up to $3.4\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$. The mean went down to $0.004\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$ during the summer months, when domestic heating and lighting fuel usage is lower. The largest fossil fuel emission estimated was located towards the north of the city, and corresponded to a crude oil refinery. Most point estimates were located on the outskirts of the city, with a few located within the central peninsula area. The road network is apparent in the figure of the prior fossil fuel fluxes displaying the corresponding transport emissions, and clearly illustrates the large contribution that road transport makes to the overall $\mathrm{CO_2}$ budget of CT.

Since we solved for weekly, rather than daily fluxes, we used a strong assumption that fossil fuel fluxes within the same week were 100% correlated. To allow the inversion to react to local conditions within a given week, no correlation was assumed between weekly fluxes. Since fossil fuel emissions were expected to be localised in space, we also assumed no spatial correlation between fossil fuel fluxes.

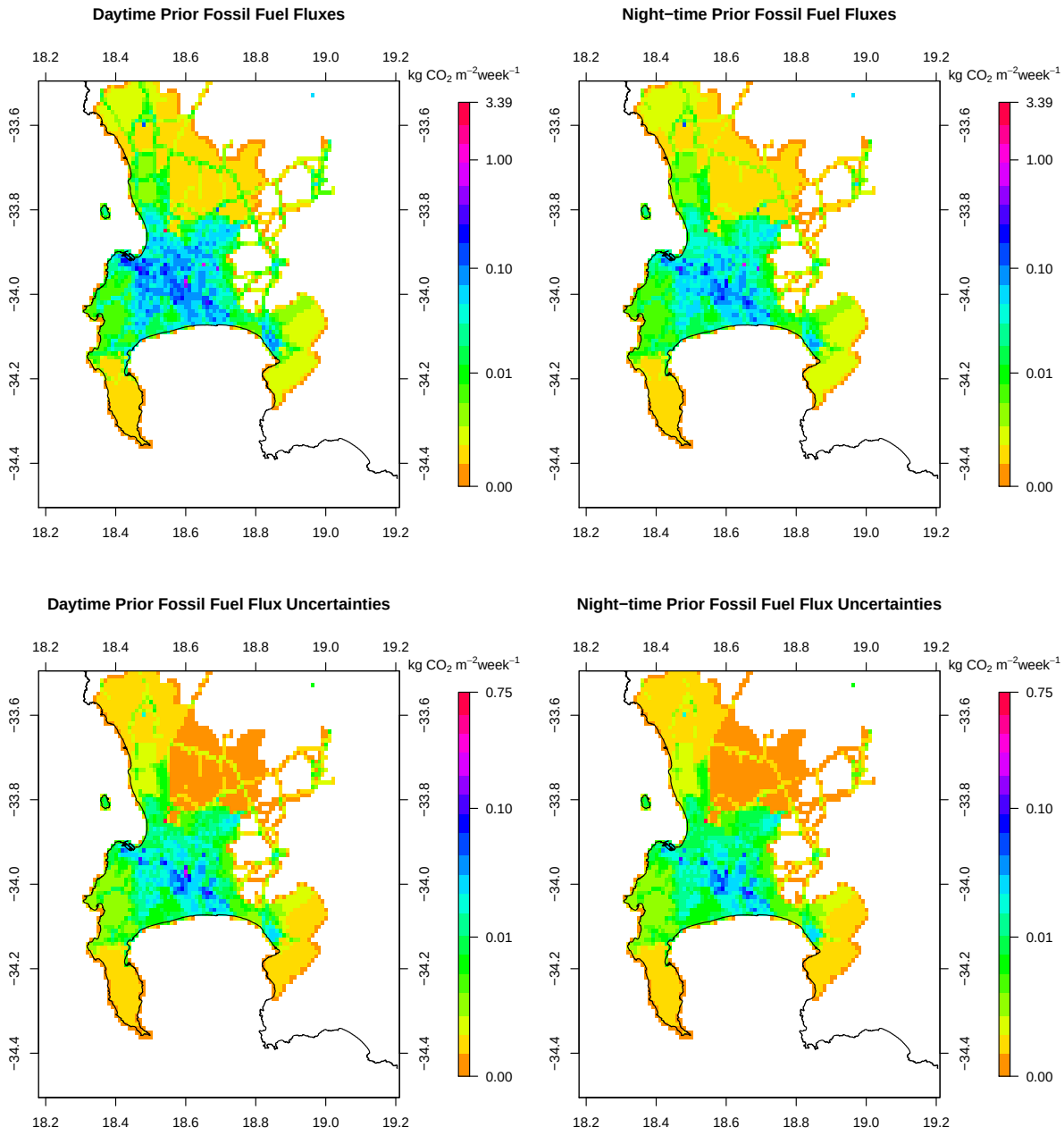


Figure 3. Prior estimates for day and night-time fossil fuel fluxes (kg CO_2 m^{-2} $week^{-1}$) and the corresponding uncertainties, expressed as standard deviations (kg CO_2 m^{-2} $week^{-1}$), for the month of March 2012. These estimates were derived from an inventory analysis for CT based on vehicle, aviation and shipping vessel count data, population census data, and fuel usage at industrial point sources. White indicates regions where the fossil fuel flux and its uncertainty are set to zero. These prior estimates are provided at a resolution of $1\,\mathrm{km} \times 1\,\mathrm{km}$ and the extent of the grid is between $34.5°$ and $33.5°$ south and between $18.2°$ and $19.2°$ east.

2.9.2 Biogenic fluxes

The uncertainty in the biogenic prior fluxes was set at the absolute value of the net primary productivity (NPP) as produced by CABLE. This is a large error relative to the prior estimate itself, but there is a great deal of uncertainty in both the productivity and respiration fluxes contributing to the NEE flux (Wang et al., 2011). The estimates of NEE are strongly dependent on the assumptions behind the model forms selected for different processes in the CABLE model. For example, the model forms used for the soil temperature-respiration function and the soil moisture-respiration function have large impacts on the NEE estimates, with resulting NEE estimates differing by over 100% compared with measurements from flux towers (Exbrayat et al., 2013). The approach of assigning either the productivity or respiration component of NEE as the uncertainty has been used by Chevallier et al. (2010). We avoided assigning a fixed proportional uncertainty to the NEE estimates as, particularly in semi-arid regions, such as those conditions found throughout South Africa, small NEE fluxes can occur as a result of both large productivity and respiration fluxes. In the CT situation, this would lead to unrealistically low estimates of the uncertainty in NEE fluxes. This is different to the approach used by Bréon et al. (2015), where an uncertainty level of 70% was assigned to biogenic fluxes, but in their case absolute NEE estimates were usually large in summer and expected to be small in winter.

To estimate covariances between the uncertainties in the NEE fluxes, we assumed an isotropic Balgovind correlation model as used in Wu et al. (2013). This helps to ensure positive-definiteness of the resulting covariance matrix. The off-diagonal covariance elements for $s_{NEE;i}$ and $s_{NEE;j}$ were calculated as:

$$C_{s_{0;NEE}}(s_{NEE;i}, s_{NEE;j}) = \sqrt{C_{s_{0;NEE}}(s_{NEE;i})}\sqrt{C_{s_{0;NEE}}(s_{NEE;j})}(1+\frac{h}{L})exp(-\frac{h}{L}) \quad (12)$$

where $s_{NEE;i}$ and $s_{NEE;j}$ are NEE fluxes in pixels *i* and *j*, $C_{s_{0;NEE}}(s_{NEE;i})$ and $C_{s_{0;NEE}}(s_{NEE;j})$ the corresponding variances in the NEE flux uncertainties in pixels *i* and *j*, the characteristic correlation length L was assumed to be $1\,\mathrm{km}$, and h is the spatial distance between pixels *i* and *j*. As for the fossil fuel fluxes, no correlation was assumed between weekly biogenic fluxes, since the inversion setup is already assuming that biogenic fluxes within the same week were 100% correlated (i.e. constant over the week).

Figure 4 shows the spatial distribution of the NEE fluxes and their uncertainties for the month of March 2012. Day-time NEE fluxes ranged between -0.19 and $0.04\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$, concentrated over areas such as the Cape Point Nature Reserve and Kogelberg Nature Reserve, located near the Hangklip lighthouse. At night the fluxes were between 0.0 and $0.06\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$. The uncertainties in the NEE daytime fluxes ranged between 0.00001 (over the ocean) and $0.30\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$, whereas at night the uncertainties ranged between between 0.000001 and $0.006\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$. Uncertainties were smaller at night because night-time biogenic activity was mainly driven by respiration, and consequently the flux estimates were smaller as well as their uncertainties. Over the full measurement period, the estimates of NEE fluxes ranged between -0.22 and 0.004 during the summer to -0.11 to $0.007\,\mathrm{kg\ CO_2\ m^{-2}\,week^{-1}}$ during mid winter.

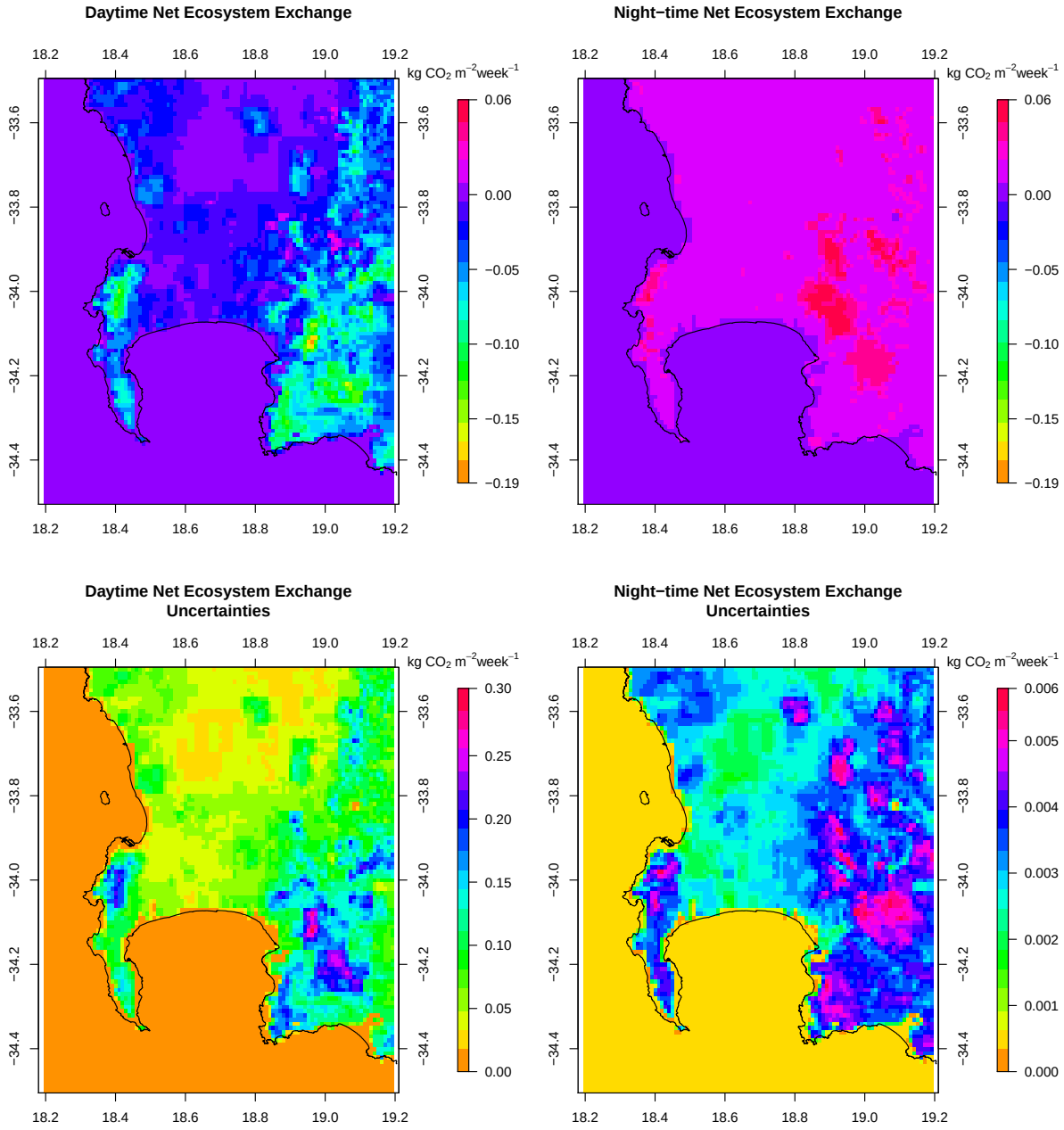


Figure 4. Prior estimates for day and night-time NEE fluxes (kg CO_2 m^{-2} $week^{-1}$) and the corresponding uncertainties, expressed as standard deviations (kg CO_2 m^{-2} $week^{-1}$), during the month of March 2012. The prior estimates were obtained from the CABLE land-atmosphere exchange model at a spatial resolution of $1\,km \times 1\,km$. The extent of the grid is between 34.5° and 33.5° south and between 18.2° and 19.2° east.

2.10 Uncertainty covariance matrix of the observations - $\mathbf{C_c}$

The uncertainties in the observations represented in $\mathbf{C_c}$ contain both the measurement error (which are known to be in the order of $0.3\,\mathrm{ppm}$) (Bréon et al., 2015; Wu et al., 2016) and the error associated with modelling the concentrations. The modelling errors result from several sources, including errors within the atmospheric transport model and aggregation errors which are due to smoothing emission estimates from localised sources within the spatial grids (Kaminski et al., 2001).

Similar to the approach adopted in the optimal network design for South Africa (Nickless et al., 2015b), an error of $2\,\mathrm{ppm}$ during the day and $4\,\mathrm{ppm}$ at night was assigned to each observation, so that night-time observations carried less weight in the inversion. These values were assigned as baseline (i.e. minimum) errors, and accounted for measurement errors, atmospheric transport modelling errors, aggregation errors and representation errors.

These errors are smaller than those for city-scale inversions conducted in the Northern Hemisphere. We justify the use of these values in our application since we are dealing with a much smaller city compared with the megacity applications, such as Paris and Indianapolis. Measurements of background $\mathrm{CO_2}$ have shown that $\mathrm{CO_2}$ concentrations in the Southern Hemisphere have smaller standard deviations. For example, for the years 2012 to 2013 the standard deviation between the monthly $\mathrm{CO_2}$ means for Mauna Loa GAW station in the Northern Hemisphere was $2.3\,\mathrm{ppm}$ (Tans and Keeling , 2016), whereas for the same time period at Cape Point the standard deviation between the monthly means was $1.6\,\mathrm{ppm}$.

We accounted for additional sources of error in the atmospheric transport model. We took into consideration that errors in the modelled $\mathrm{CO_2}$ concentrations due to the transport model would be larger when the wind speed was lower (Bréon et al., 2015), and this would be compounded at night when the planetary boundary layer height was lower and less stable (Feng et al., 2016). Additional error ranging between 0 and $1\,\mathrm{ppm}$ was added to the daytime uncertainty of $2\,\mathrm{ppm}$, linearly scaled depending on the wind speed, with $0\,\mathrm{ppm}$ added when wind speeds were high ($20\,\mathrm{m\,s^{-1}}$) and $1\,\mathrm{ppm}$ added when the wind speed was close to zero. At night the additional uncertainty ranged between 0 and $4\,\mathrm{ppm}$.

We also considered the standard deviation of the measured $\mathrm{CO_2}$ concentrations during each hour. It would be expected that if there was a large amount of variability between the instantaneous measurements at the site, that the atmospheric transport model would be more likely to make errors during this period. The variance of the observed $\mathrm{CO_2}$ concentrations that contributed towards the mean estimate of the $\mathrm{CO_2}$ concentration for that hour, was added to the overall uncertainty. Therefore each hour had a customised observation error dependant on the prevailing conditions at the measurement site. Therefore the total observation error for hour *k*, as a variance, is given as:

$$C_c(k,k) = {C_{c;base}}^2 + {C_{c;wind}}^2 + {C_{c;obs}}^2 \tag{13}$$

where $C_{c;base}$ is the baseline observation error of $2\,\mathrm{ppm}$ during the day and $4\,\mathrm{ppm}$ during the night, $C_{c;wind}$ is the additional error due to the wind speed conditions which ranged between 0 and 1, and $C_{c;obs}$ is the standard deviation of the observed concentrations within that hour. A time series of the customised observation errors is provided in Figure 5. The final observation errors could reach up to 10 or $15\,\mathrm{ppm}$ at night, reducing the weight of these measurements in the estimation of the prior fluxes.

Temporal correlation between the observation errors was accounted for in an analogous manner to which covariance terms were estimated for the NEE flux uncertainties. The characteristic correlation length L was assumed to be $1\,\mathrm{hour}$, and h was the temporal distance between observations.

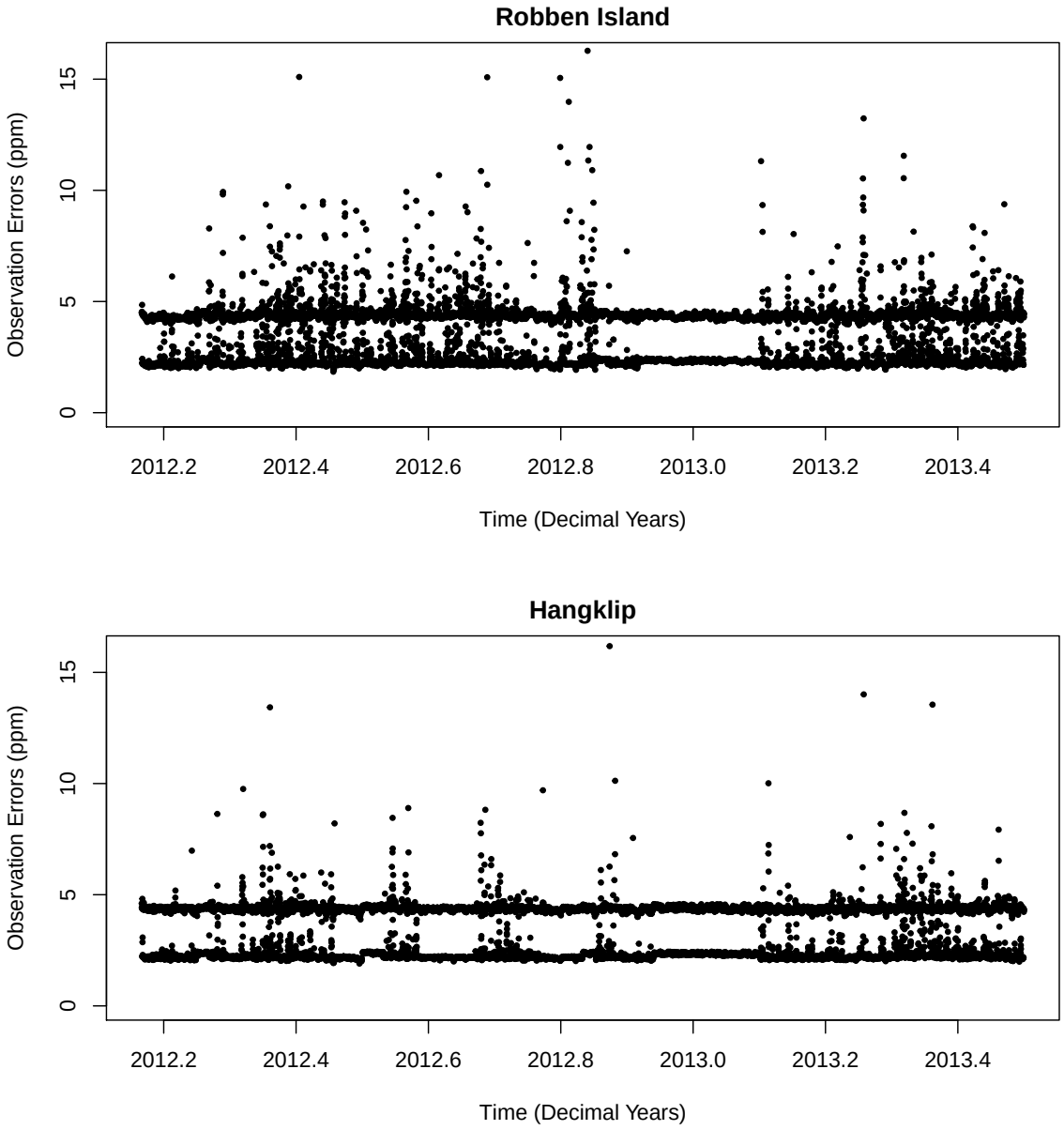


Figure 5. Time series of the customised observation errors (ppm) assigned to the CO_2 concentration measurement for each hour at the Robben Island and Hangklip measurement sites. The errors consist of a baseline error (set as $2\,\mathrm{ppm}$ during the day and $4\,\mathrm{ppm}$ at night), and additional atmospheric model errors based on prevailing wind speed and the variation in the instantaneous CO_2 observations within an hour. The two distinct sets of points for each site arises due to the night-time observation errors set to be larger than daytime observation errors.

2.11 Model Assessment

In order to assess the appropriateness of the uncertainty covariance matrices $\mathbf{C_c}$ and $\mathbf{C_{s_0}}$, the χ^2 statistic, as described in Tarantola (2005) can be employed to determine the minimum value of the statistic:

$$\chi_1^2 = \frac{1}{\nu}(\mathbf{Hs_0} - \mathbf{c})^T(\mathbf{HC_{s_0}H}^T + \mathbf{C_c})^{-1}(\mathbf{Hs_0} - \mathbf{c}) \tag{14}$$

where ν is the dimension of the data space, in this case the length of observations in the inversion.

The squared residuals from the inversion (squared differences between observed and modelled concentrations) should follow the χ^2 distribution with degrees of freedom equal to the number of observations (Michalak et al., 2005; Tarantola, 2005). Dividing this statistic by the degrees of freedom should yield a χ_1^2 distribution. Values lower than one indicate that the uncertainty is too large, and values greater than one indicate that the uncertainty prescribed is lower than it should be. The error in the assignment of the uncertainty could be in either $\mathbf{C_c}$ or $\mathbf{C_{s_0}}$ (or both).

Sensitivity analyses carried out on the specification of the covariance matrices have indicated that these errors are most likely contained in $\mathbf{C_{s_0}}$. These analyses are presented in a companion paper. In order to ensure the suitability of $\mathbf{C_{s_0}}$, the prior variances were multiplied by a factor of two. This ensured that the χ_1^2 statistic was close to a value of one for almost all months of the inversion. A single scaling factor was used to adjust all the prior flux variances. An alternative to a single value scaling factor will be considered in a subsequent paper.

Using the χ^2 statistic to scale or estimate covariance parameters has been implemented by Lauvaux et al. (2016) and Michalak et al. (2005). Lauvaux et al. (2016) used the χ^2 statistic to scale the elements of the observation error covariance matrix. An alternative to manually scaling the elements of either $\mathbf{C_{s_0}}$ or $\mathbf{C_c}$, is to use a hierarchical Bayes approach to estimate hyper-parameters for the covariance matrix, which are estimated based on the observed concentrations (Ganesan et al., 2014).

3 Results

In this paper we concentrate on the results of the reference inversion, as described in the previous section. We present sensitivity analyses elsewhere. Additional information on the distribution and time series of the observed concentrations at Robben Island, Hangklip and Cape Point over the 16 month period are provided in supplementary material section 1.1. Information is also provided on the assessment of the goodness-of-fit of the prescribed covariance structures in supplementary material section 1.2 to justify the use of the scaling factor of 2 to increase the original estimates of the uncertainties in the prior fluxes to get the χ^2 statistic closer to one. The average wind speed and direction, supplied as monthly maps of the wind fields across the domain, as modelled by CCAM are provided in section 1.3.

3.1 Modelled Concentrations

The time series of the prior and posterior modelled concentrations at Robben Island were compared with the observed concentrations (Figures 6). The prior estimates tended to be in the correct range for CO_2 concentration measurements, but could be higher or lower compared with the observations by as much as $100\,\mathrm{ppm}$. It is possible to test whether our assumed uncertainties in the prior fluxes are consistent with the misfit between the prior modelled concentrations and observations. Michalak et al. (2005) pointed out that the covariances of the differences between the prior simulation and observations is given by the matrix $\mathbf{HC_{s_0}H}^T + \mathbf{C_c}$. This matrix accounts for both the uncertainty in the prior fluxes and in the observations. The square root of the diagonal elements of this matrix had a similar distribution to the absolute mismatches between the observations and prior modelled concentrations, showing that the set-up is statistically consistent. The prior concentrations tended to spike at the same time as the observations, but these spikes were usually larger in the prior modelled concentrations.

As one would expect, the agreement between the posterior modelled concentrations and the observations was much stronger compared with the prior estimates. The posterior concentrations appeared to track the observed concentrations during localised "pollution" events. For example, in March to April 2012 all except one of the spikes in the observed CO_2 concentration was replicated in the posterior concentrations. The agreement can be assessed by means of the intraclass correlation coefficient (ICC) (Shrout and Fleiss, 1979), which is a stronger condition than correlation. Values close to zero indicate poor agreement while values close to one indicate strong agreement. The ICC was low at 0.03 (95% CI: 0.01 to 0.06), but still significant, between the observed and prior modelled concentrations, but went up to 0.59 (95% CI: 0.57 to 0.61) between the observed and posterior modelled concentrations.

We define prior residuals as the difference between the observed and prior modelled concentrations, and posterior residuals as the difference between the observed and posterior modelled concentrations. A time series plot of the prior and posterior residuals given in Figure 6 indicates more clearly how large the misfits between the modelled and observed concentrations can get. The prior residuals could be large in either the positive or negative direction, up to $100\,\mathrm{ppm}$ and occasionally out by as much as $200\,\mathrm{ppm}$. The posterior residuals were much closer to the zero line, with the highest deviation equal to $33\,\mathrm{ppm}$. The bias in the prior modelled concentrations was -2.9 ppm. The standard deviation of the prior residuals was $21.4\,\mathrm{ppm}$ (interquartile range between -9.1 and $3.7\,\mathrm{ppm}$), indicating a large amount of spread in the residuals. The bias in the posterior modelled

concentrations went down to 0.5 ppm and the standard deviation of residuals reduced to 3.9 ppm (interquartile range -1.5 and 1.5 ppm), showing a significant reduction in the misfit compared with the prior modelled concentrations. Compared with the standard deviation of the observed concentrations, which was 5.02 ppm, the standard deviation of the posterior residuals was lower by 1.1 ppm, indicating that the uncertainty in the posterior estimates of the concentrations was well below the expected variability around the observed concentrations.

The time series of the observed, prior and posterior concentrations at Hangklip reveal a similar result compared with those for Robben Island (Figure 7). The prior estimates could be much larger or smaller compared with the observed concentrations. The posterior concentration estimates matched much more closely with the observed concentrations compared with those for Robben Island. The ICC between the observed and prior modelled concentrations was similar to Robben Island at 0.03 (95% CI: 0.003 to 0.05), but the agreement between the observed and posterior modelled concentrations was better with an ICC of 0.76 (95% CI: 0.75 to 0.77). The prior residuals at the Hangklip site tended to be less extreme compared with those for Robben Island, with a maximum deviation of 117 ppm in either direction (Figure 7). The summary statistics of the residuals indicate that the mean bias in the prior estimates was 2.4 ppm with standard deviation equal to 17.6 (interquartile range between -2.3 and 6.5 ppm). For the posterior residuals, the bias was reduced to 0.04 ppm with standard deviation equal to 2.46 (interquartile range -1.1 to 0.8), the standard deviation lower by 1.4 ppm compared with the standard deviation of the observed concentrations, which was 3.89 ppm.

The observed and modelled concentrations and their misfits are provided separately for day and night concentrations in Figures 6 and 7. There is no notable difference in the degree of misfit between day and night at either site. The large improvement in the representativeness of the posterior concentrations in relation to the observed concentrations at both sites lends confidence to the reference inversion's ability to adjust the estimates of the fluxes to better match the true fluxes in the region.

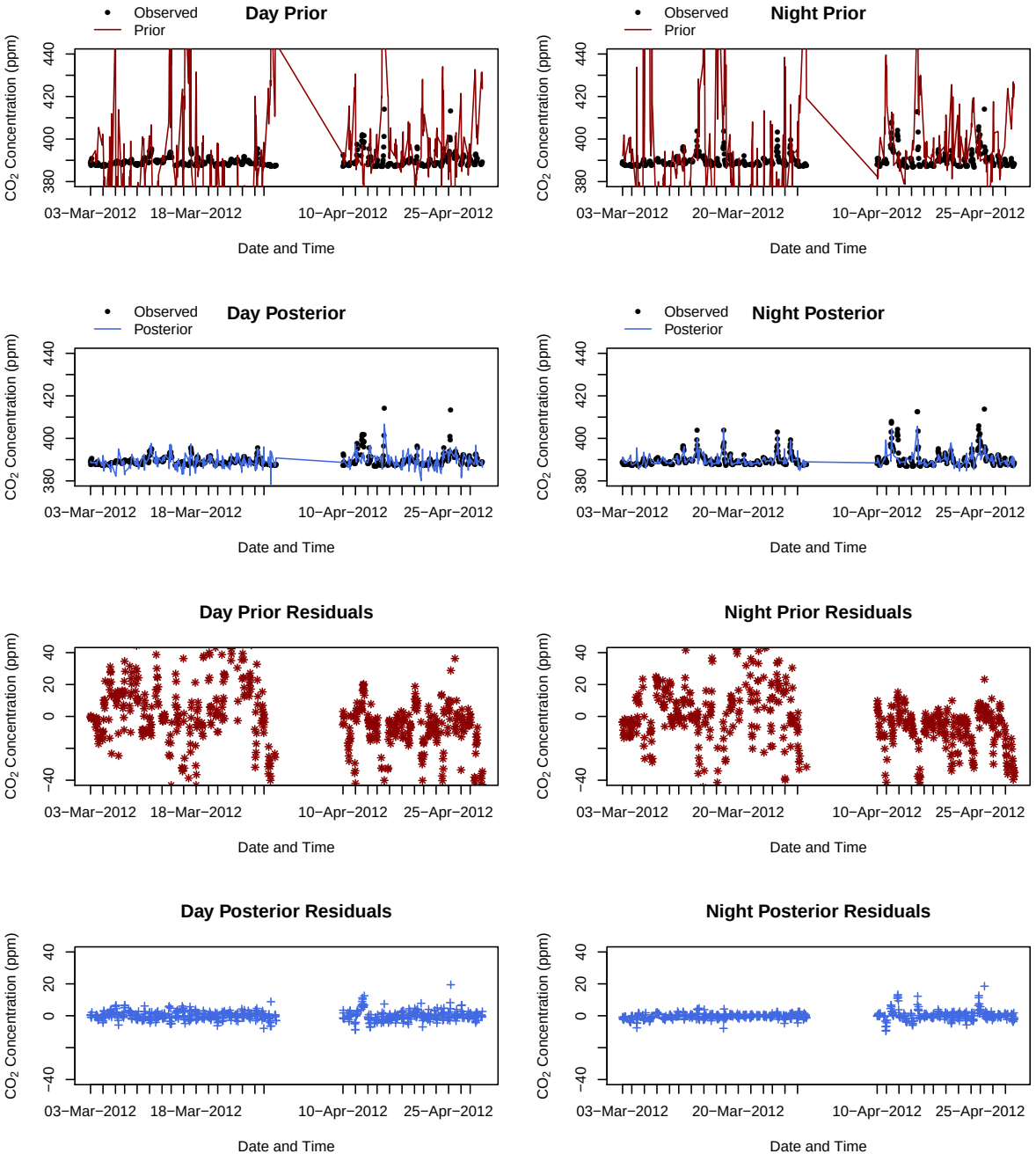


Figure 6. The top 4 panels provide a time series of the observed, prior and posterior modelled concentrations at the Robben Island site. The time series is separated into day and night-time periods. The residuals between the observed and prior/posterior modelled concentrations, defined as the difference between the observed and modelled concentrations, are provided in the lower panel 4 panels. The first two months are presented here and remainder of the time series is presented in the supplementary material section 1.4.

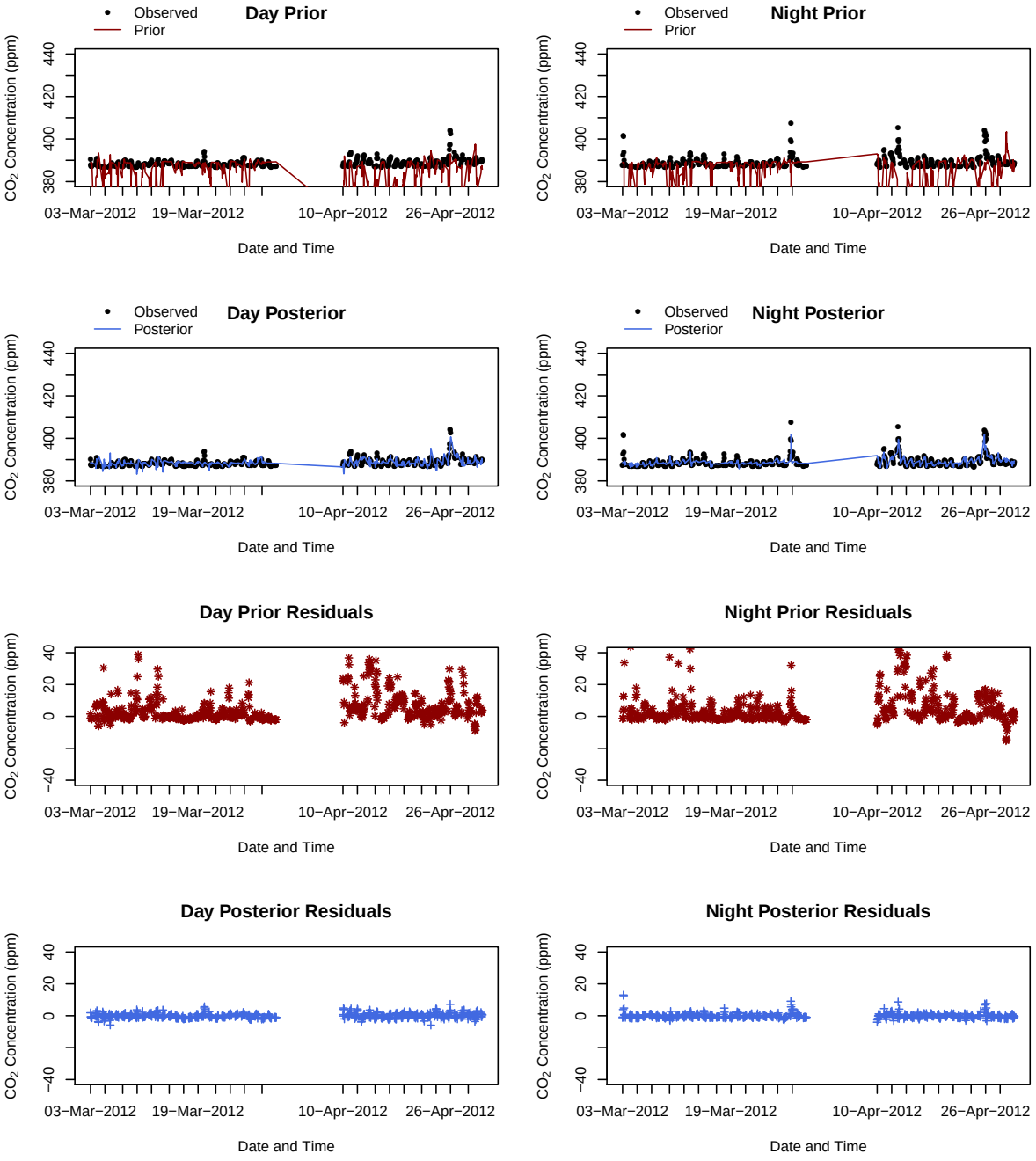


Figure 7. The top 4 panels provide a time series of the observed, prior and posterior modelled concentrations at the Hangklip site. The time series is separated into day and night-time periods. The residuals between the observed and prior/posterior modelled concentrations, defined as the difference between the observed and modelled concentrations, are provided in the lower panel 4 panels. The first two months are presented here and remainder of the time series is presented in the supplementary material section 1.4

The mean working week and weekend diurnal cycles in the observed, prior and posterior modelled concentrations are shown for each site and for each month in the supplementary material (supplementary material section 1.5). Figure 8 provides the mean working week and weekend diurnal cycle over the full measurement period. For Robben Island, the mean concentrations for each hour indicate that the emissions are overestimated by the prior estimates. The posterior modelled concentrations are much closer to the observed concentrations, replicating the peak in concentrations to be between 8:00 and 9:00 in the morning and the trough in concentrations to occur between 15:00 and 18:00. Overall the cycle in the posterior concentrations is flatter compared with that of the observed concentrations. The observed concentrations during the week are usually slightly higher compared with those over the weekend. The posterior estimates show a smaller deviation between the week and weekend concentrations at each hour of the day, particularly around mid-morning, compared with the observed week and weekend concentrations.

The prior estimates for the Hangklip measurement show the opposite bias compared with Robben Island, with prior modelled concentrations lower at each hour compared with the observed concentrations. The posterior modelled concentrations for Hangklip overlap closely with the observed concentrations. When compared with Robben Island, there is slightly less separation between the working week and weekend concentrations at each hour. This should be expected as the concentrations observed at the Hangklip site are more dominated by biogenic sources compared with Robben Island. The closest fossil fuel sources are those from transport and domestic emissions. The main road through this area carries a large amount of commercial traffic during the week, and over the weekend the area is frequented by weekend residents and tourists, and therefore anthropogenic activity is not expected to be much lower over the weekend. The posterior concentrations show that the inversion was able to replicate the separation between the mean hourly working week and weekend concentrations shown by the observed concentrations for most hours of the day, particularly the difference between these concentrations occurring between the hours of 7:00 and 9:00.

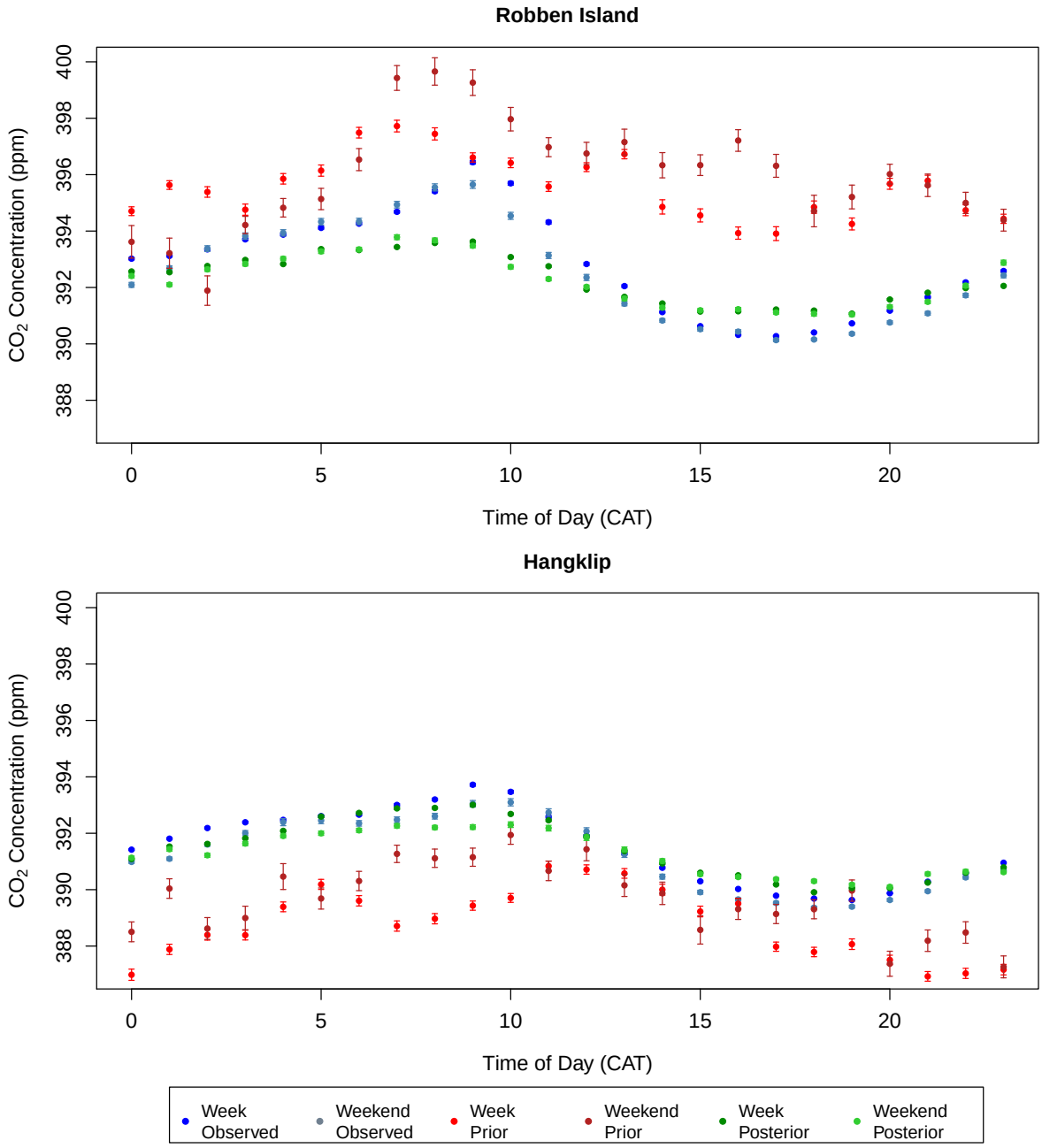


Figure 8. The hourly diurnal cycle (mean concentrations for each hour with 95% confidence interval) in the observed, prior and posterior modelled CO_2 concentrations (ppm) over the full measurement period from March 2012 until June 2013, separated by working week and weekend, and plotted separately for Robben Island (top) and Hangklip (bottom) measurement sites. The diurnal plots are separated into working week and weekend observed concentrations (blue and light blue), working week and weekend prior modelled concentrations (red and dark red), and working week and weekend posterior modelled concentrations (green and light green).

The contributions by the fossil fuel and NEE fluxes to the modelled concentrations were determined. These are displayed in Figures 9 and 10 for Robben Island and Hangklip respectively for March to June 2012. July 2012 to June 2013 are supplied in the supplementary material section 1.6. The prior contributions by these two fluxes were of similar magnitude at both sites, but with the fossil fuel fluxes increasing CO_2 concentrations and the NEE fluxes reducing the concentration. After the inversion, the contributions of the posterior NEE fluxes to the modelled concentrations were much more modified by the inversion compared with those from the posterior fossil fuel fluxes. Moreover, the adjustments made to the NEE fluxes resulted in the contributions to the CO_2 concentrations that were much more similar in magnitude to the fossil fuel fluxes, to the extent that the uptake of CO_2 due to biogenic processes cancelled out the contributions made by the fossil fuel fluxes.

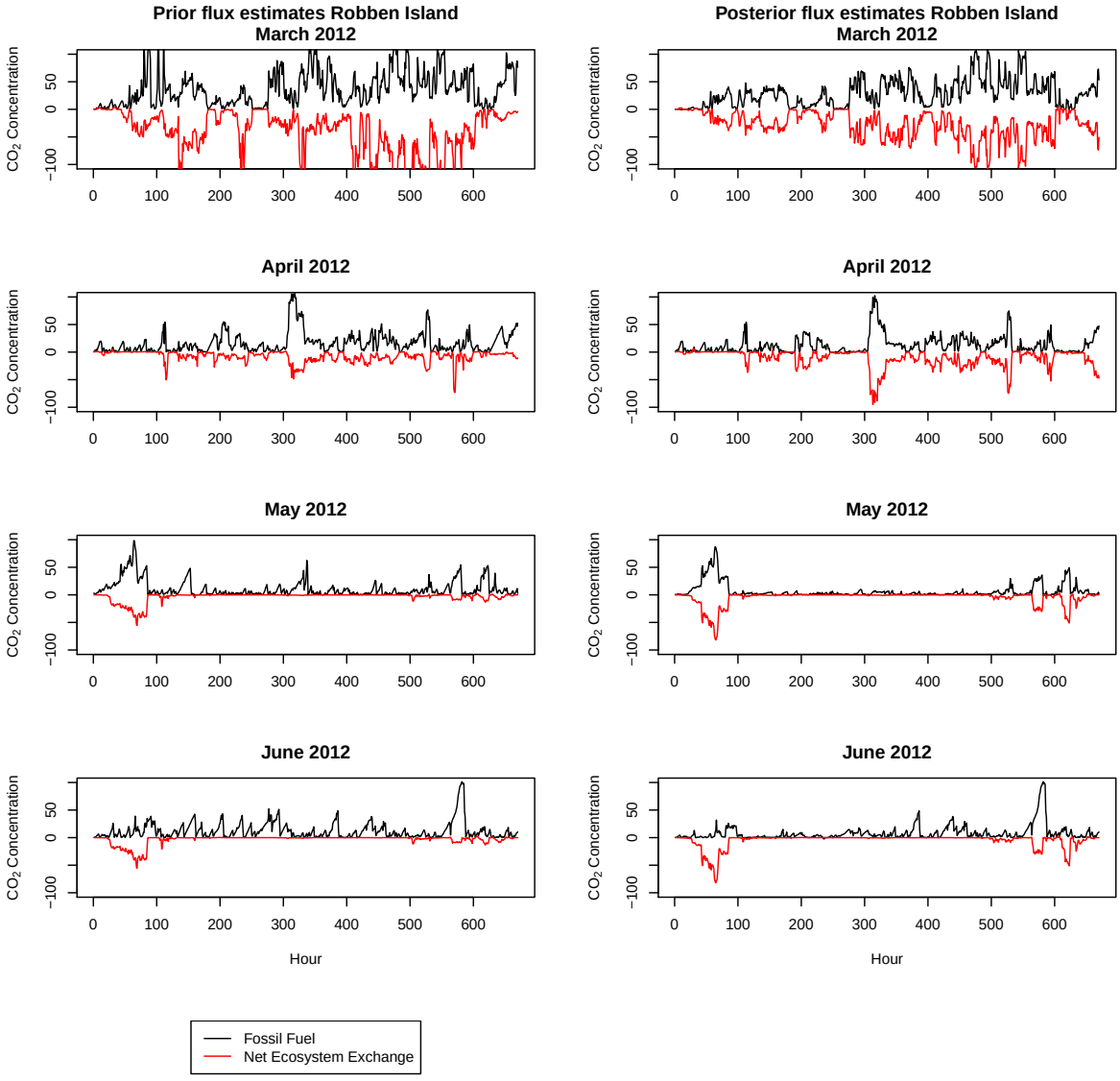


Figure 9. Prior and posterior contributions of the fossil fuel and NEE surface fluxes to the modelled CO_2 concentrations (ppm) at Robben Island from March 2012 until June 2012.

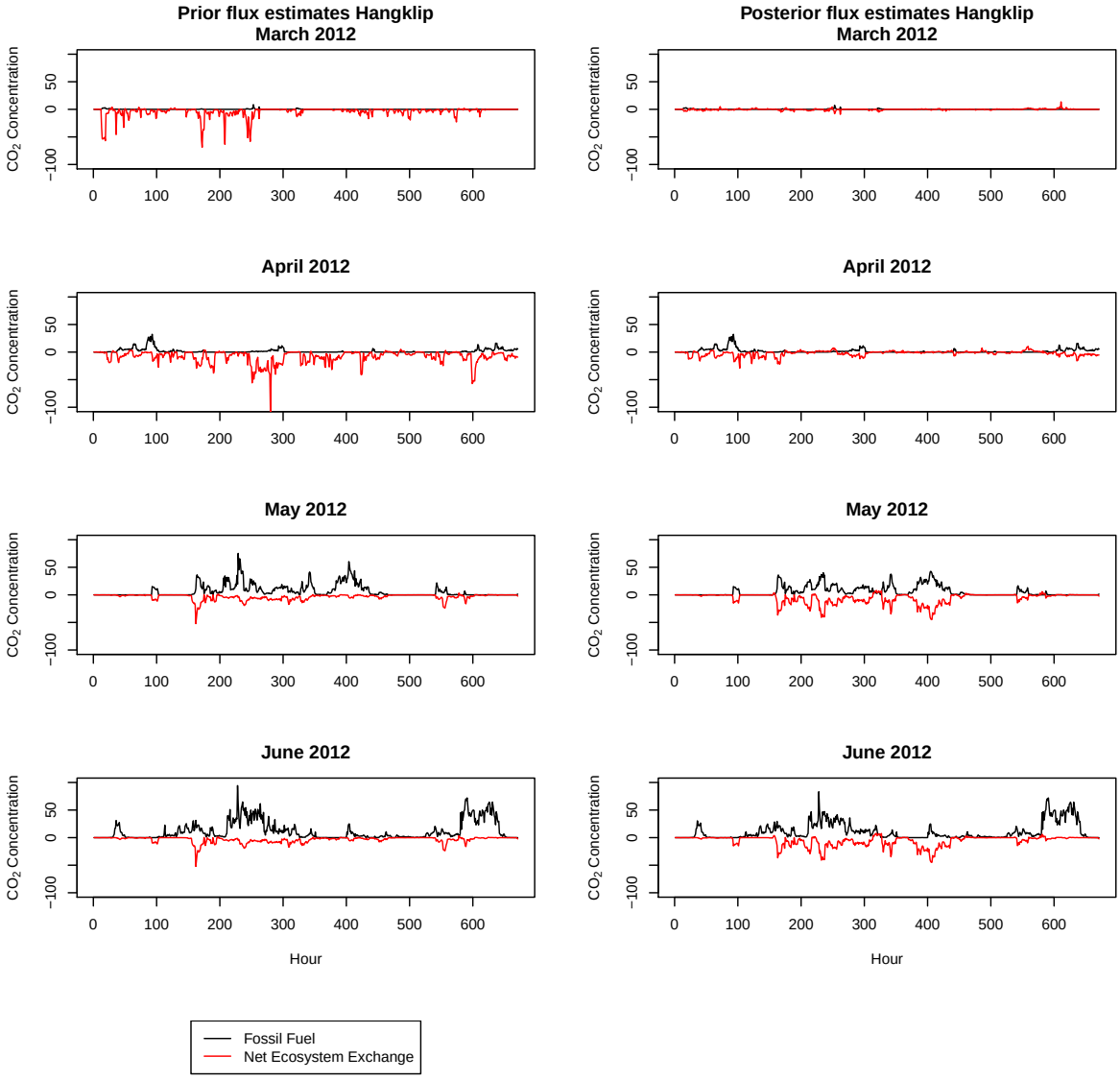


Figure 10. Prior and posterior contributions of the fossil fuel and NEE surface fluxes to the modelled CO_2 concentrations (ppm) at Hangklip from March 2012 until June 2012.

The inversion-corrected CO_2 concentrations at the domain boundary are obtained from the posterior source estimates. Therefore these concentrations are extracted from the solution of $\mathbf{s}$. Figure 11 provides a time series of the prior and posterior concentration estimates. The northern and eastern domain borders are terrestrial, whereas the ocean borders the south and west. For all four domain boundaries, across the total measurement period, the inversion has made only small innovations, with the posterior estimates remaining within the 95% confidence limits of the prior concentrations. Only the northern and eastern terrestrial boundaries showed some deviations from the priors between May and June 2012, and between March and April 2013. As these are the terrestrial boundaries of the domain, the concentrations here would be subject to influences from outside of the domain, and it would therefore be expected that the inversion would need to provide greater adjustments to these boundary concentrations. The ocean boundaries would be expected to have concentrations close to the background concentrations provided by the Cape Point measurement site.

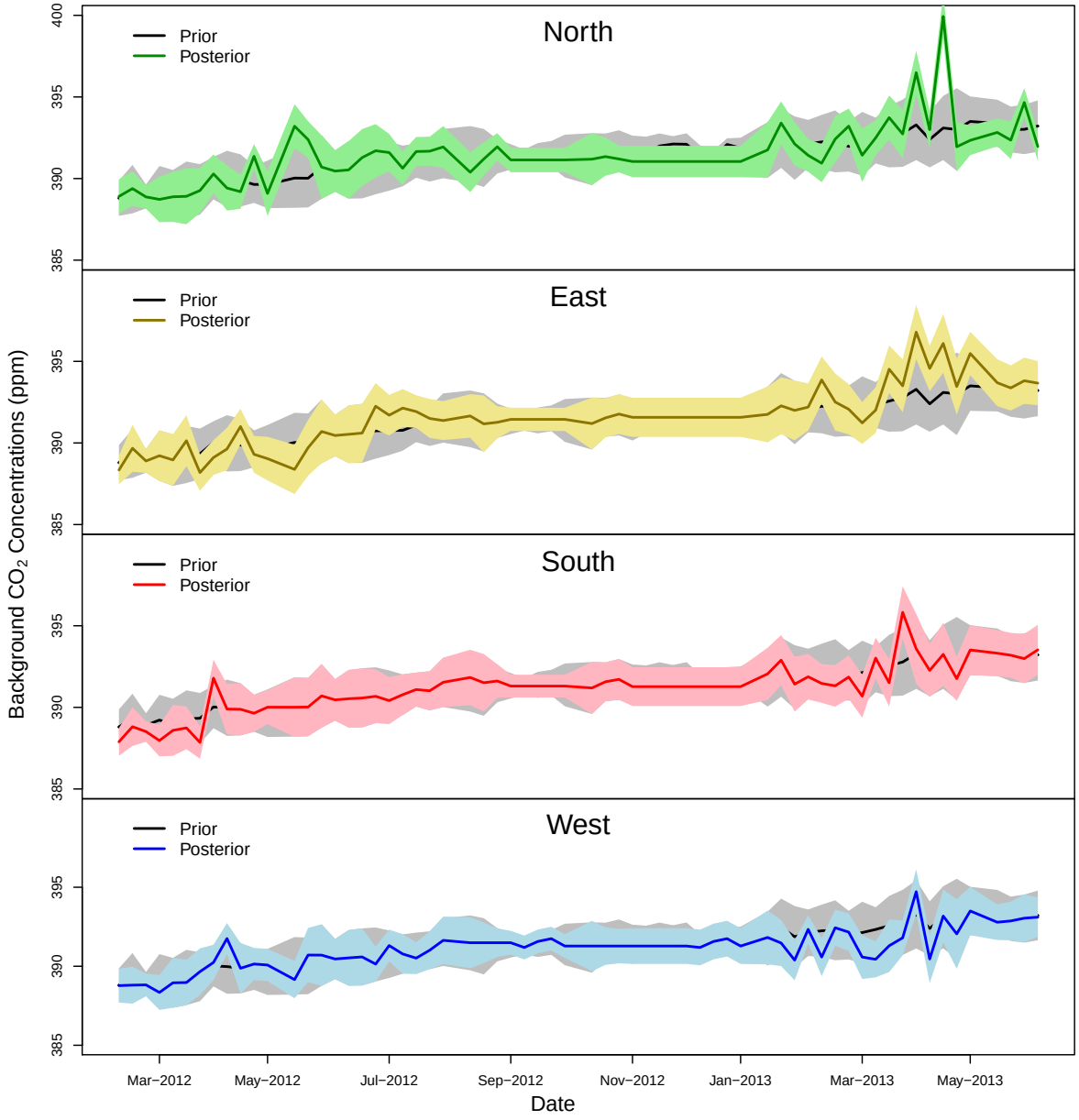


Figure 11. Time series with 95% confidence interval (represented by the shaded area) of the prior (black line) and posterior estimates of the CO_2 concentrations (ppm) at the domain boundaries (north - green, east - yellow, south - red, west - blue). The prior estimates are the same for each cardinal direction, and are obtained from the Cape Point percentile-filtered observations. The posterior estimates for the concentrations are solved for as additional unknowns in the reference inversion.

3.2 Weekly flux Estimates

We refer to the difference between the prior and posterior flux estimates as the innovations. The impact of the inversion on the flux estimates can be assessed through the size and direction of these innovations and through the reduction in the flux uncertainties. Figure 12 shows the innovations in the total flux estimates for each pixel in $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ for the month of May 2012, as well as the percentage reduction in the flux uncertainty, the percentage reduction in the fossil fuel emissions and change in NEE fluxes. The mean total weekly flux of a pixel was obtained by first deriving the total flux from the six fossil fuel and NEE flux estimates for that week, and then taking the mean of the four weekly fluxes obtained for the month. The innovations for the total flux were calculated as the difference between the total prior and total posterior weekly fluxes. Positive innovations indicate that the prior fluxes were too far in the positive direction, and that the flux should be adjusted towards the negative direction, whereas negative innovations indicate that prior fluxes were too far in the negative direction and should be adjusted towards the positive direction. The uncertainties for the total flux in each pixel are represented by standard deviations which have been derived from the elements of the prior and posterior uncertainty covariance matrices of the fluxes. The figure displays the percentage uncertainty reduction within each pixel. The percentage reductions in the total fossil fuel flux are presented, rather than the absolute changes from the prior to posterior flux estimates in each pixel as these changes were small and the inversion generally resulted in a reduction in the fossil fuel flux estimate. The inversion had more freedom to make changes to the NEE estimates as the uncertainties prescribed were large relative to the prior NEE flux estimates, and therefore the absolute changes from prior to posterior are presented in $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$.

May falls within the winter rainfall season of the Western Cape region. The innovations indicate that the total flux for the pixel over the petrol refinery, which had the largest prior flux estimate, was overestimated by the prior ($9.43\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) relative to the posterior estimate ($6.62\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) by an amount of $2.81\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ (Fig. 12). Innovations were generally small ranging between -0.001 and $0.003\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$. The area around the natural reserves, such as Cape Point and Kogelberg Nature Reserves, had innovations by the inversion that were close to zero or slightly negative, indicating that the amount of carbon uptake (i.e. negative NEE flux) in these regions was overestimated by the CABLE model. The largest negative innovation by the inversion was -$0.08\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, from -0.03 up to $0.04\,\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, over a pixel in the Cape Point Nature Reserve. The prior flux indicated the pixel to be a slight carbon sink of $\mathrm{CO_2}$ whereas the inversion adjusted the flux to be a slight $\mathrm{CO_2}$ source. With respect to the rest of the domain, excluding the crude oil refinery, the most substantial innovations were made over the central business district (CBD) area to the south east of Robben Island. These innovations were positive, indicating that the fossil fuel fluxes were overestimated by the prior estimates.

The percentage change in the fossil fuel estimates, from prior to posterior, indicate the changes tended to be small, except on Robben Island itself, where percentage changes were up to 75% indicating that the emissions on the island were significantly reduced by the inversion (Fig. 12). There were a few pixels which had negative change, indicating that the inversion increased fossil fuel emissions, just north of the island. Located on the north eastern shore of Robben Island is a diesel-fuelled power generation plant, as well as desalination plant which is powered by this station. Increases in the fossil fuel fluxes may be due to emissions arising from these activities which have not been accounted for in the inventory analysis. In the inventory analysis

there was no fuel information available for any industrial sources on Robben Island, but fossil fuel emissions were included due to domestic and transport activities, therefore these could have been adjusted by the inversion. The inventory analysis does not take into account explicitly the shipping routes going into CT harbour, or into Robben Island harbour, but rather all the emissions are concentrated within CT harbour, where the shipping information is available. This could also result in the inversion adjusting emissions on the island to deal with near shipping related emissions. There was a region in the Western Cape interior to the east of Robben Island which had slightly increased fossil fuel emissions. The inversion had the effect of mainly reducing fossil fuel emissions along the south east transect extending from Robben Island over the CBD towards Hangklip, or leaving the emission unchanged.

The inversion had a much larger impact on the terrestrial NEE fluxes (Fig. 12). This is unsurprising, as the relative uncertainties of the prior NEE fluxes were much larger compared with those of the prior fossil fuel fluxes. The area of the domain experiencing innovation from the inversion was also much more widespread compared with the innovations made to the fossil fuel fluxes. This is in part due to the correlation specified between the NEE fluxes of adjacent pixels, but not between their fossil fuel fluxes. The majority of the innovation over the domain was close to zero, between -0.02 and 0.02 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, indicating that the inversion was making small absolute adjustments to the NEE flux estimates. Over the CBD region, the adjustments were the largest, up to 0.32 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ and these differences were positive indicating that the posterior fluxes were more in the negative direction and therefore the inversion was acting to reduce total emissions of $\mathrm{CO_2}$ over the CBD region relative to the prior estimates, through changes to the NEE fluxes. The natural region around Cape Point and within the Kogelberg Nature Reserve north of Hangklip showed slight negative changes in the NEE fluxes, indicating that the uptake of $\mathrm{CO_2}$ was reduced by the inversion making the total flux in these regions more positive towards the atmosphere.

In terms of the percentage standard deviation reduction (Fig. 12), the largest reductions occurred over the natural areas, particularly Cape Point to the south of Robben Island, where the posterior uncertainty over the area was significantly lower, by over 50%, compared with the prior uncertainties. Significant reductions are also shown over largely agricultural areas to the north of the CBD region. Over the CBD area itself, the reductions were present, reaching values of close to 60% over a few central CBD pixels, but generally smaller compared with the uncertainty reductions over natural areas in the domain which reached levels as high as 92%.

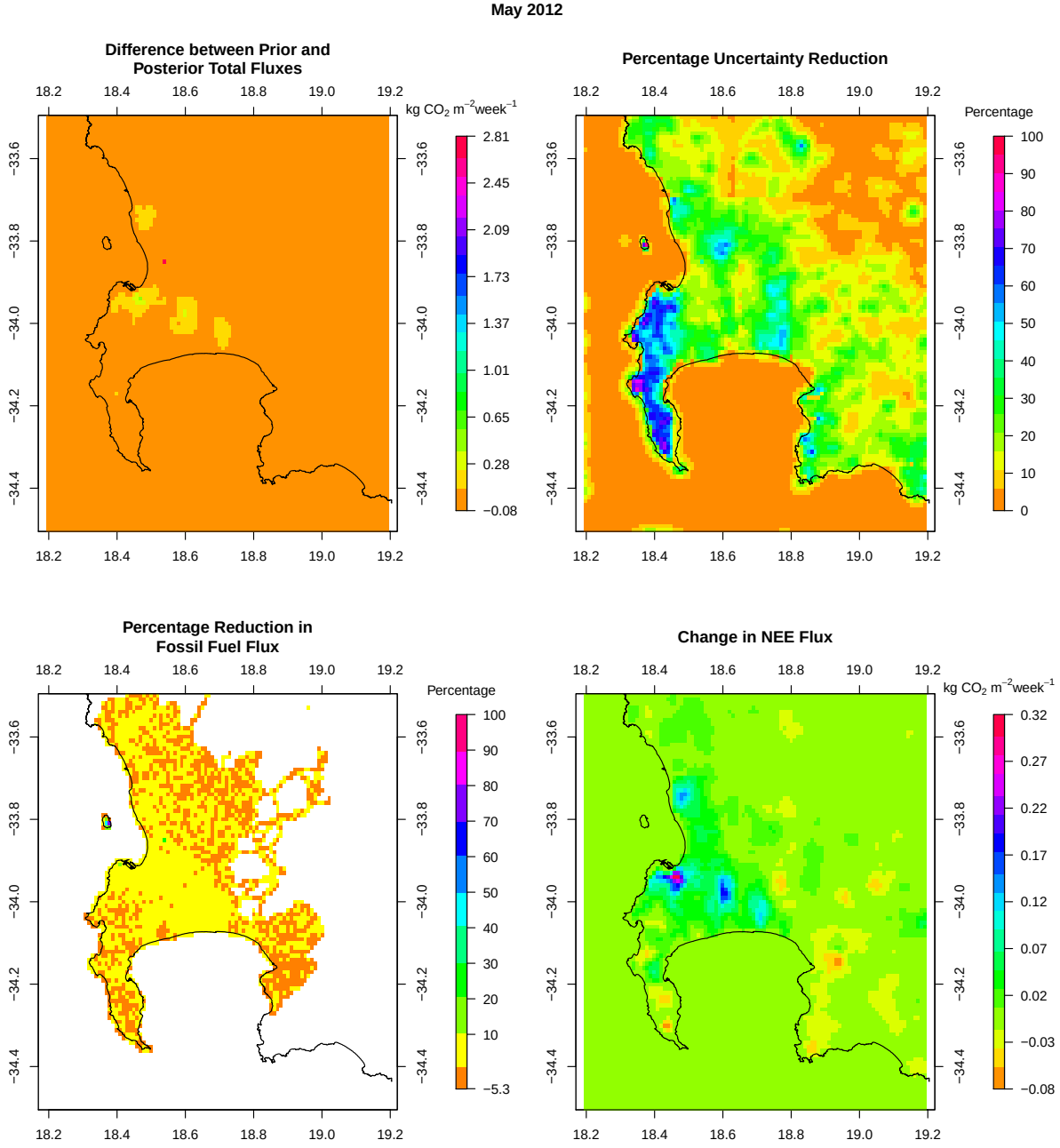


Figure 12. (Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for May 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating posterior CO_2 fluxes were made more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

September is the beginning of the summer months in the Western Cape region, when temperatures start to rise, and the mean monthly rainfall reduces. The innovations across the region during this month were dominated by negative values with the majority ranging between -0.03 and -0.01 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, indicating that the inversion increased emissions of $\mathrm{CO_2}$ over the majority of the domain. The maximum increase in $\mathrm{CO_2}$ fluxes of 0.2 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ occurred over an agricultural region north east of the CBD region, from -0.10 up to 0.08 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, where mainly vineyards are located (Fig. 13). The inversion increased the total fluxes by an amount close to 0.2 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ over a further five areas. One of these areas, also over agricultural land, lies further inland to the east of the area with largest increase in $\mathrm{CO_2}$ fluxes. Three more regions lie to the north, north east and east of the Hangklip site. These three regions are all within the Hottentots-Holland Mountain catchment area which is largely dominated by vineyard agriculture. The fifth area is located within the Kogelberg Nature Reserve near the Hangklip site. The maximum reduction of 2.1 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ occurred over the crude oil refinery site, from 9.4 down to 7.2 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ (Fig. 13). The inversion made small reductions to the fluxes near the CBD area of CT, or left the posterior fluxes to be similar to the priors.

Innovations made to the fossil fuel fluxes were mainly made on the transect of the city running between Robben Island and Hangklip, as well as to fossil fuel emissions on Robben Island itself, similar to the month of May. The maximum percentage adjustment to the fossil fuel fluxes was 51.1%, and the mean innovation close to zero, with almost all innovations positive, indicating that the posterior estimates were smaller relative to the priors. Robben Island itself showed a mix of positive and negative innovations, with posterior fluxes larger than the priors on the west of the island but smaller than the priors on the east of the island.

The effect of the inversion on the NEE fluxes was to make these fluxes more negative over the CBD region, and made $\mathrm{CO_2}$ fluxes more positive relative to the priors over the Table Mountain region, over the agricultural area to the north of the CBD, and over the natural regions near Hangklip. The fluxes over the CBD were made more negative by up to 0.21 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ through adjustments to the NEE fluxes. The areas to the north of the domain which had prior NEE estimates between -0.20 and -0.10 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$ were made more positive, with posterior estimates ranging between -0.10 and 0.10 $\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$, where the largest changes were over the vineyard agricultural areas. The natural area surrounding Hangklip site showed negative innovations between the prior and posterior fluxes, indicating that the inversion adjusted the negative prior NEE fluxes by making these more positive. Therefore the CABLE model appears to have overestimated the amount of $\mathrm{CO_2}$ uptake during this period.

Most reductions in the total flux uncertainty ranged between 2.3 and 18.6%, with a maximum reduction of 88.5% (Fig. 13). The largest uncertainty reductions induced by the inversion occurred over the natural areas bordering on the CBD, particularly over the Table Mountain National Park, as well as the natural areas surrounding the Hangklip site. The areas to the east of Robben Island over the Durbanville and Bellville townships, comprised of a mix of residential suburbs, vineyard agricultural areas and industrial areas, also showed reductions in the uncertainties of the fluxes.

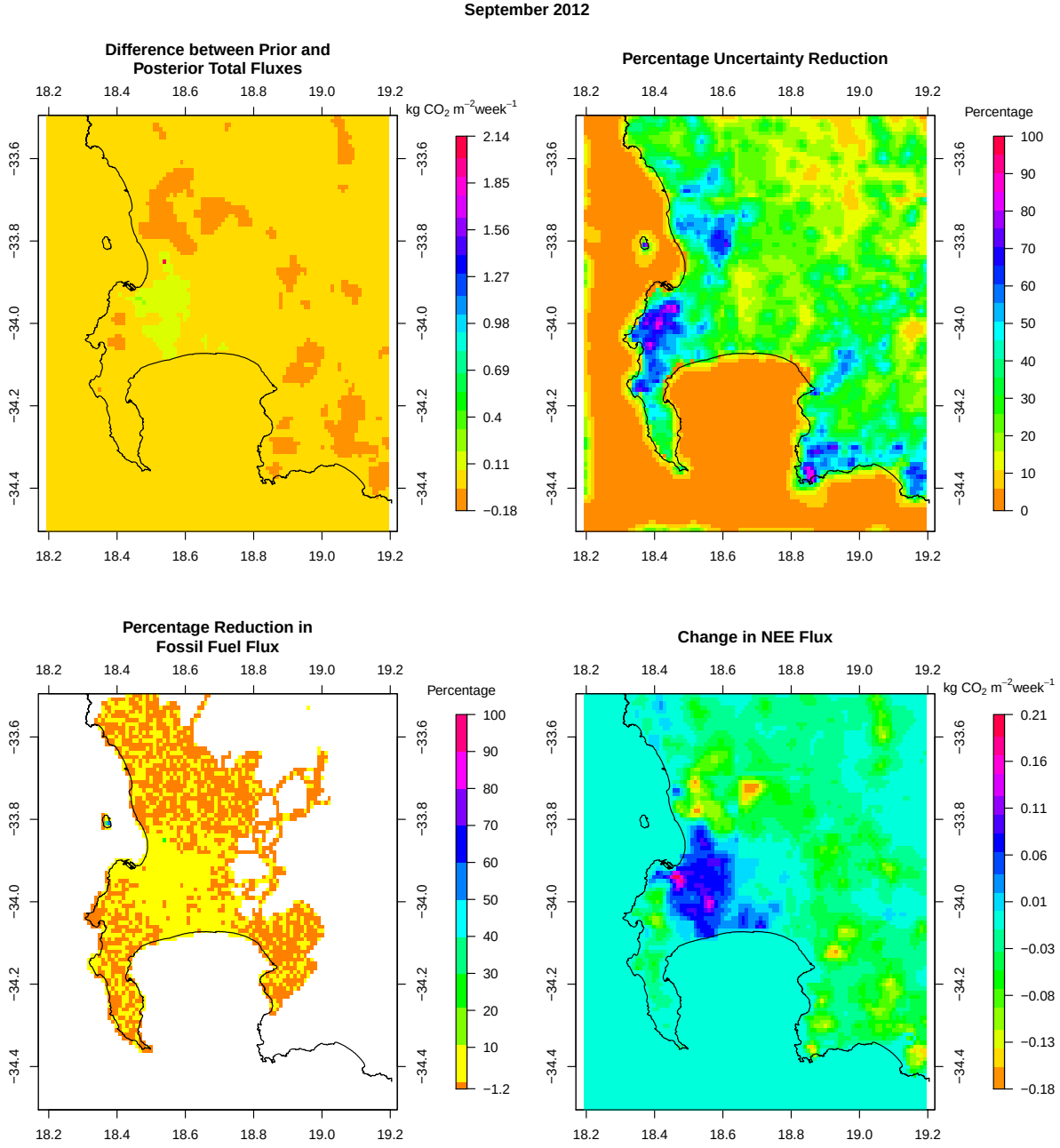


Figure 13. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for September 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior $\mathrm{CO_2}$ fluxes were made more positive by the inversion compared with the prior estimates. Extent: between $34.5°$ and $33.5°$ south and between $18.2°$ and $19.2°$ east.

Figure 14 presents boxplots of the pixel-level weekly fluxes over the domain for each month, and the supplementary material (supplementary material section 1.7) displays the spatial extent of the adjustments made to prior flux estimates by the inversion, as well as the uncertainty reductions, for each month. The limits of the range for the posterior pixel-level weekly fluxes are more negative compared with those of the prior estimates, indicating that for all the months, the inversion tended to reduce the emission of CO_2. This is also evident from the maps of differences in total flux estimates between the prior and posterior estimates. Specifically, the inversion tended to reduce the fossil fuel emissions, evident from the shift downwards in the distribution of the posterior pixel-level fossil fuel fluxes compared with the prior fluxes. The variability in the posterior fluxes across pixels was slightly reduced compared with the variability in the prior fluxes. The pixel-level NEE fluxes, although generally smaller in magnitude and range, were the most altered by the inversion, where these changes were related to the season in which the month fell. For most months, the overall mean in the NEE fluxes became more positive, indicating less uptake of CO_2 than predicted by the CABLE model. But for the months of June and July, which occur mid-winter, the mean NEE was made more negative. The minimum values were also at least twice the minimum value of the prior estimates for all months, and this is also evident from the maps of change in NEE from prior to posterior, which show that the inversion reduced the CO_2 flux towards the atmosphere over the CBD region by altering the NEE fluxes within those pixels.

For most months the percentage uncertainty reduction for the pixel-level total flux reached over 90% for at least some of the pixels, with a maximum uncertainty reduction over a pixel of 97.7% in March 2012. The lowest maximum reduction in the uncertainty of the total flux in a pixel in a month was 68.2% in June 2012.

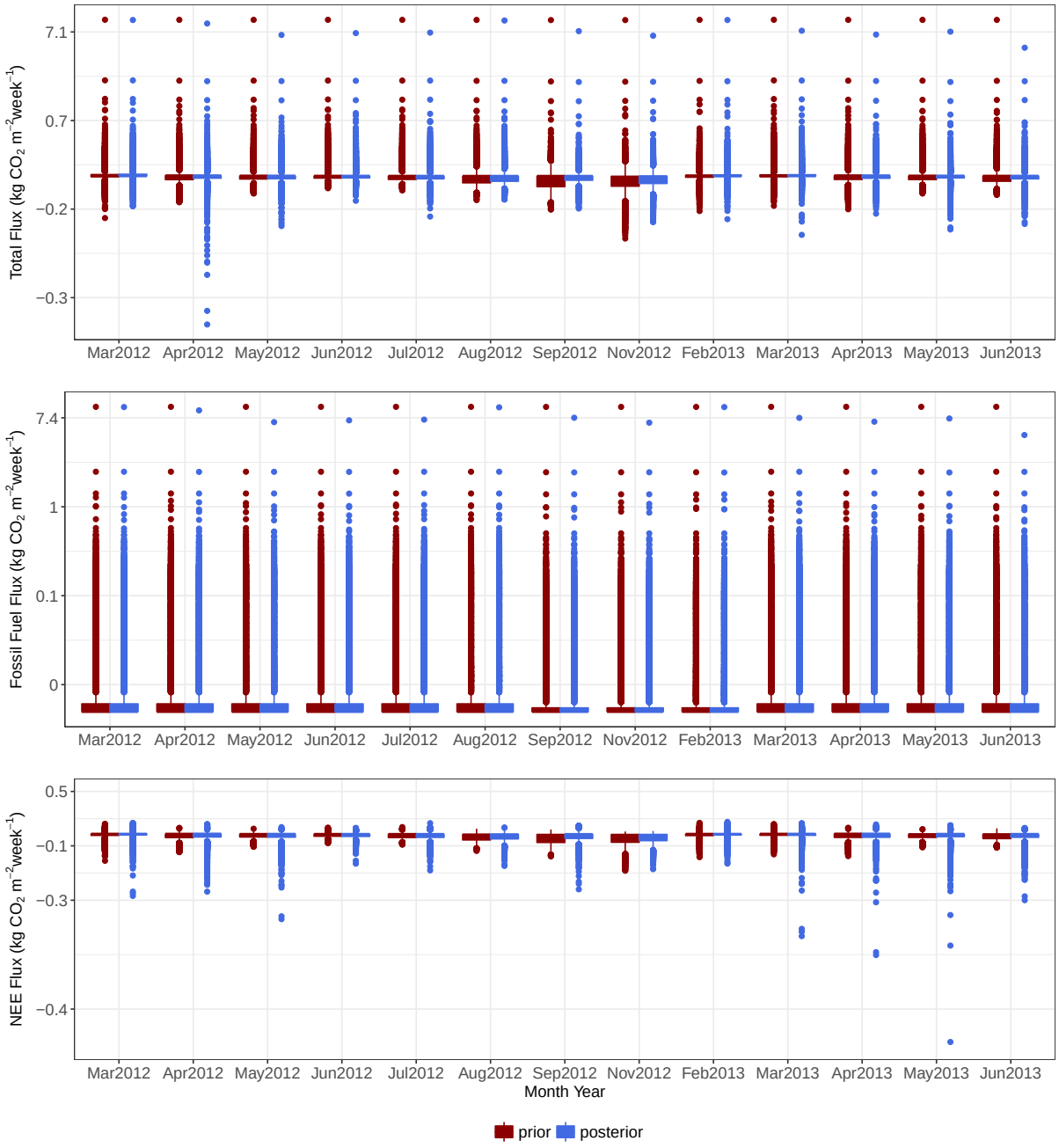


Figure 14. Boxplots of the pixel-level weekly prior and posterior flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for each month. These plots present summary statistics calculated over all pixels in the domain of the pixel-level mean weekly fluxes. The y-axis is presented on a log scale.

An example pixel, located near the CBD sources, was selected in order to investigate posterior uncertainty spatial covariances in the fluxes resulting from the inversion. For a given week, the total CO_2 flux is composed of six sources: working week and weekend fossil fuel sources, both day and night, and the day and night biogenic sources; and each of these sources could have a non-zero covariance term between itself and the same source but from surrounding pixels, or with one of the other five sources from the same pixel or from surrounding pixels.

The posterior uncertainty covariances between the daytime working week flux of the selected pixel and other sources are only notably different from zero for working week daytime fossil fuel fluxes, working week night-time fossil fuel fluxes, and the daytime biogenic fluxes. The uncertainty covariances between the daytime working week fluxes reveal that non-zero covariances do not necessarily have to be close in proximity to the selected pixel, and negative and positive covariances can cluster in space. The uncertainty covariances with the night-time fossil fuel fluxes were larger than those during the day, but were limited to a few pixels close to the selected pixel. These covariances ranged between -0.15 and 0.09 $\mathrm{g}^2\ \mathrm{CO_2\,m^{-4}\,week^{-2}}$. The non-zero uncertainty covariances with NEE fluxes were larger (between -1.50 and 0.88 $\mathrm{g}^2\ \mathrm{CO_2\,m^{-4}\,week^{-2}}$) and fluctuated between patches of positive and negative values. Closer to the CBD area there was a distinct region of positive uncertainty covariance between the fossil fuel source of the selected pixel and the NEE fluxes from a region over the Table Mountain area and a negative covariance patch south of the CBD. The eastern terrestrial part of the domain had patches of positive and negative covariances. When converting these covariances are into correlations, correlations are small; no bigger than 0.001 in either direction.

The covariances between fossil fuel and NEE flux uncertainties are small because the uncertainties in the prior modelled concentrations that are attributed to the flux contributions ($\mathbf{HC_{s_0}H}^T$) are small relative to the uncertainties specified for the modelled concentration errors ($\mathbf{C_c}$). This is not because our prior uncertainty is small but because the transport Jacobian only projects fluxes from individual pixels weakly into modelled concentrations. As the uncertainty in the modelled concentration errors is decreased, the size of the posterior off-diagonal covariance elements between the fossil fuel and NEE flux uncertainties from the same pixel increases. This can easily be confirmed through the use of a toy inversion system using typical values for $\mathbf{H}$, $\mathbf{C_{s_0}}$ and $\mathbf{C_c}$ from our inversion system. This is presented in section 1.8 of the supplementary material.

The sum of the covariances in the uncertainties between the selected fossil fuel flux and all other fluxes equals -25.8 $\mathrm{g}^2\ \mathrm{CO_2\,m^{-4}\,week^{-2}}$. Therefore the covariances associated with this pixel's fossil fuel flux would reduce the variance associated with the total pixel-level flux by 51.7, where the total uncertainty of the fossil fuel flux, as a variance, was 233.7 $\mathrm{g}^2\ \mathrm{CO_2\,m^{-4}\,week^{-2}}$.

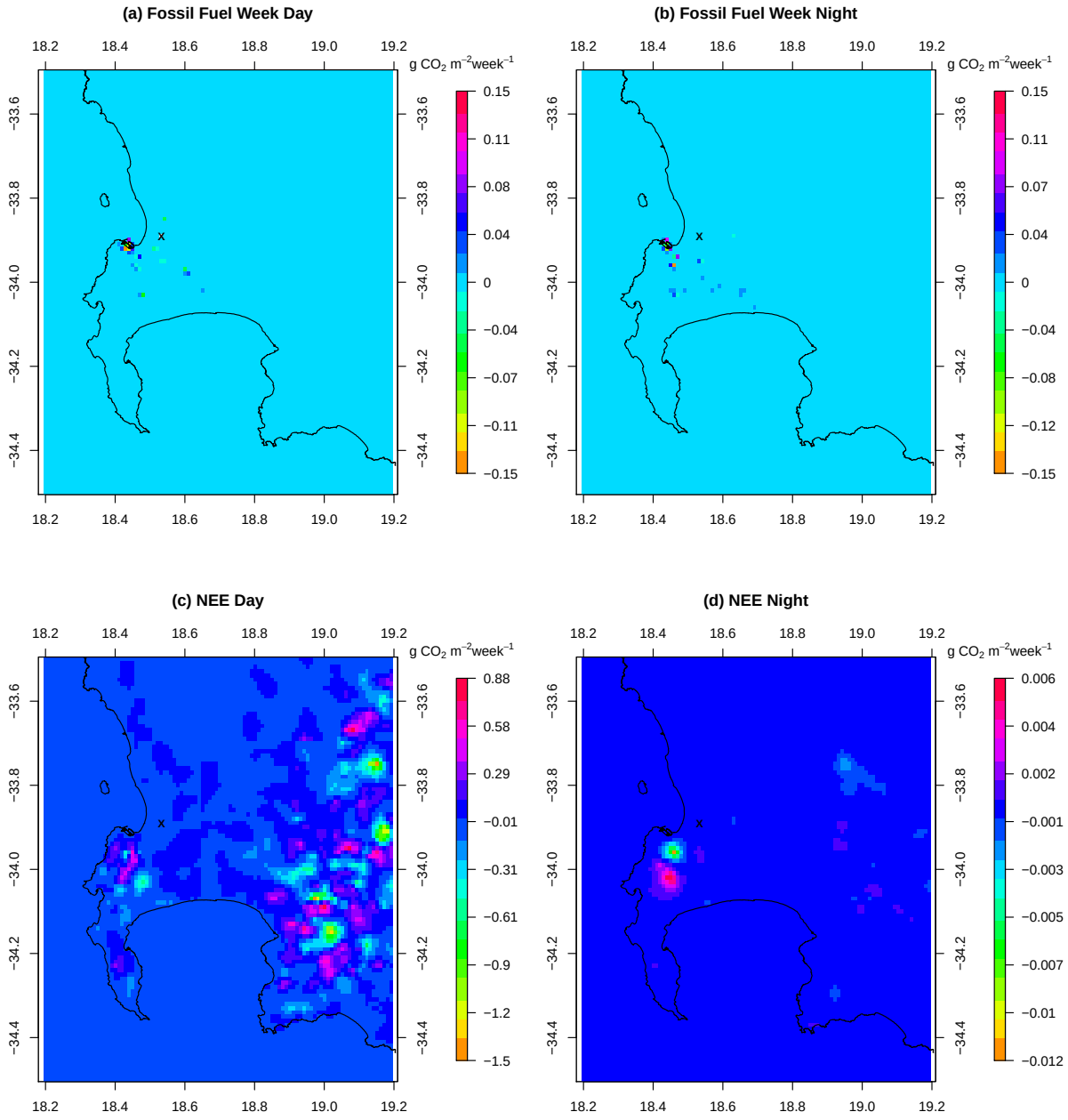


Figure 15. Posterior covariances ($g^2\ CO_2\ m^{-4}\ week^{-2}$) between the uncertainty in the fossil fuel working week daytime flux during the first week of March 2012 in the selected pixel (marked by X) and (a) all other fossil fuel working week daytime fluxes; (b) fossil fuel working week night-time fluxes within this pixel and all other pixels; (c) NEE daytime fluxes within this pixel and all other pixels; (d) NEE night-time fluxes within this pixel and all other pixels. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

3.3 Weekly Totals

Three time series plots of the total weekly estimates of CO_2 flux over the full spatial domain (i.e. all fluxes aggregated over all pixels) are presented in Figure 16. The total flux estimate for a week represents the sum of all the fossil fuel and NEE fluxes for that week. For the prior total flux uncertainties, the only non-zero covariances are those imposed between the uncertainties in NEE fluxes of neighbouring pixels. These positive covariance terms increase the total uncertainty. The uncertainties in the posterior total fluxes includes the posterior covariance terms between the uncertainties in the fossil fuel and NEE fluxes, which are generally negative and have the overall effect when summed of reducing the uncertainty in the total flux estimate.

Uncertainties in the prior total fossil fuel fluxes ranged between 1.3 and 1.5 $\mathrm{kt\ CO_2\ week^{-1}}$, whereas the posterior uncertainties ranged between 0.9 and 1.5 $\mathrm{kt\ CO_2\ week^{-1}}$. For the NEE fluxes the uncertainties ranged between 23.6 and 57.3 $\mathrm{kt\ CO_2\ week^{-1}}$ before the inversion and between 15.8 and 47.1 $\mathrm{kt\ CO_2\ week^{-1}}$ after the inversion. Uncertainties in the total fluxes were similar to those of the NEE fluxes, with prior uncertainties ranging between 23.6 and 57.3 and posterior uncertainties ranging between 15.8 and 47.1 $\mathrm{kt\ CO_2\ week^{-1}}$. The median percentage uncertainty reduction in the total weekly flux was 28.0 % and ranged between 2.3 and 50.5 %, with the largest reduction occurring in March 2012.

The posterior total estimate for the emission of CO_2 from the domain was within the confidence bounds of the prior total estimate for the majority of the period from March 2012 until June 2013. The confidence bounds of the posterior total estimates were narrower compared with those of the prior total estimates. Total prior flux estimates ranged between 139.5 and -386.8 $\mathrm{kt\ CO_2\ week^{-1}}$, with the maximum total during March 2013 and the minimum total flux occurring in November 2012. The posterior totals ranged between 149.5 and -375.1 $\mathrm{kt\ CO_2\ week^{-1}}$, with the maximum occurring in March 2013 and the minimum in October 2012. During the winter months, from March to July, the posterior fluxes fell within the uncertainty limits of the prior fluxes. The posterior total flux moved outside of the prior's confidence limits during August and September 2012, which was during the South African spring period. Posterior estimates were larger compared with the prior estimates. Data were missing during October 2012 and from December 2012 with January 2013, and therefore the estimates are completely overlapping during these periods. When observations were available during the summer months in November 2012 and February 2013 the posterior fluxes were more positive compared with the prior estimates.

The total prior fossil fuel flux was flat and close to 150 $\mathrm{kt\ CO_2\ week^{-1}}$ during the winter months, and close to 110 $\mathrm{kt\ CO_2\ week^{-1}}$ during the summer months. This stepped effect in the fossil fuel fluxes is due to the simple representation of the domestic emissions in the fossil fuel inventory. It is unlikely that fossil fuel emissions would have a sharp change between summer and winter. As a separate sensitivity analysis presented in a companion paper we adjusted the assumption of domestic emissions such that domestic emissions were distributed uniformly during the year. The inversion had the effect of reducing the total estimate, particularly during May 2012 and between March 2013 and June 2013, to a value of as low as 138 $\mathrm{kt\ CO_2\ week^{-1}}$ during the winter months and to a value of 107 $\mathrm{kt\ CO_2\ week^{-1}}$ during the summer months and early winter 2013. The posterior total fossil fluxes and confidence bounds for June 2013 were outside of those for the prior estimates for the full month. Compared with the total flux, the range of the total fossil fuel fluxes was much narrower (between 100 and 160 $\mathrm{kt\ CO_2\ week^{-1}}$), and the confidence bounds around the estimates were also narrower. This is not immediately apparent

from the plot, but the range of the y-axis needed to be adjusted for the fossil fuel fluxes, otherwise it would appear as a thin line if plotted on the same range as the total fluxes.

Total prior NEE fluxes ranged between values close to zero and -494.9 $\mathrm{kt\ CO_2\ week^{-1}}$ and between zero and -483.1 $\mathrm{kt\ CO_2\ week^{-1}}$ for the posterior estimates. During the winter months posterior estimates were generally contained within the limits of the prior estimates, except for May 2012, where the total NEE flux was slightly lower compared with the prior. From August 2012 to September 2012, the posterior total NEE fluxes were well above the total prior estimates, indicating that the total uptake of $\mathrm{CO_2}$ by the domain was reduced by the inversion during this period.

Comparisons of the NEE and fossil fuel fluxes to the total estimates show that the variability in the total flux estimates was driven by variability in the NEE fluxes, and differences between the posterior and prior total estimates were mainly driven by adjustments to the NEE fluxes induced by inversion. As would be expected, due to the large uncertainty around the NEE prior fluxes, the inversion was not strongly constrained by these priors, and therefore mainly adjusted the NEE fluxes, rather than the fossil fuel fluxes, so that the modelled concentrations better matched the observed concentrations. This is unsurprising as we provided no information regarding what proportion of the observed $\mathrm{CO_2}$ concentration was attributable to fossil fuel contributions and what proportion was due to NEE fluxes.

Temperatures are at their highest between January and March. This is also the region's dry period. Both the prior and posterior NEE fluxes were most positive during this period, with NEE fluxes close to zero. The posterior NEE fluxes were more positive than the prior fluxes during these periods, i.e. the posterior estimates indicated less uptake of $\mathrm{CO_2}$ by biogenic processes compared with the prior estimates. The minimum temperatures occurred between July and October, and during this period the NEE estimates were negative, indicating $\mathrm{CO_2}$ uptake. The peak in the uptake by biogenic processes occurred in September to early November. This was during the spring period, after the winter rainfall period, and when the temperatures in the region begin to increase. The posterior estimates were less negative than the prior NEE fluxes, indicating that the NEE fluxes estimated by CABLE may have overestimated the amount of $\mathrm{CO_2}$ uptake.

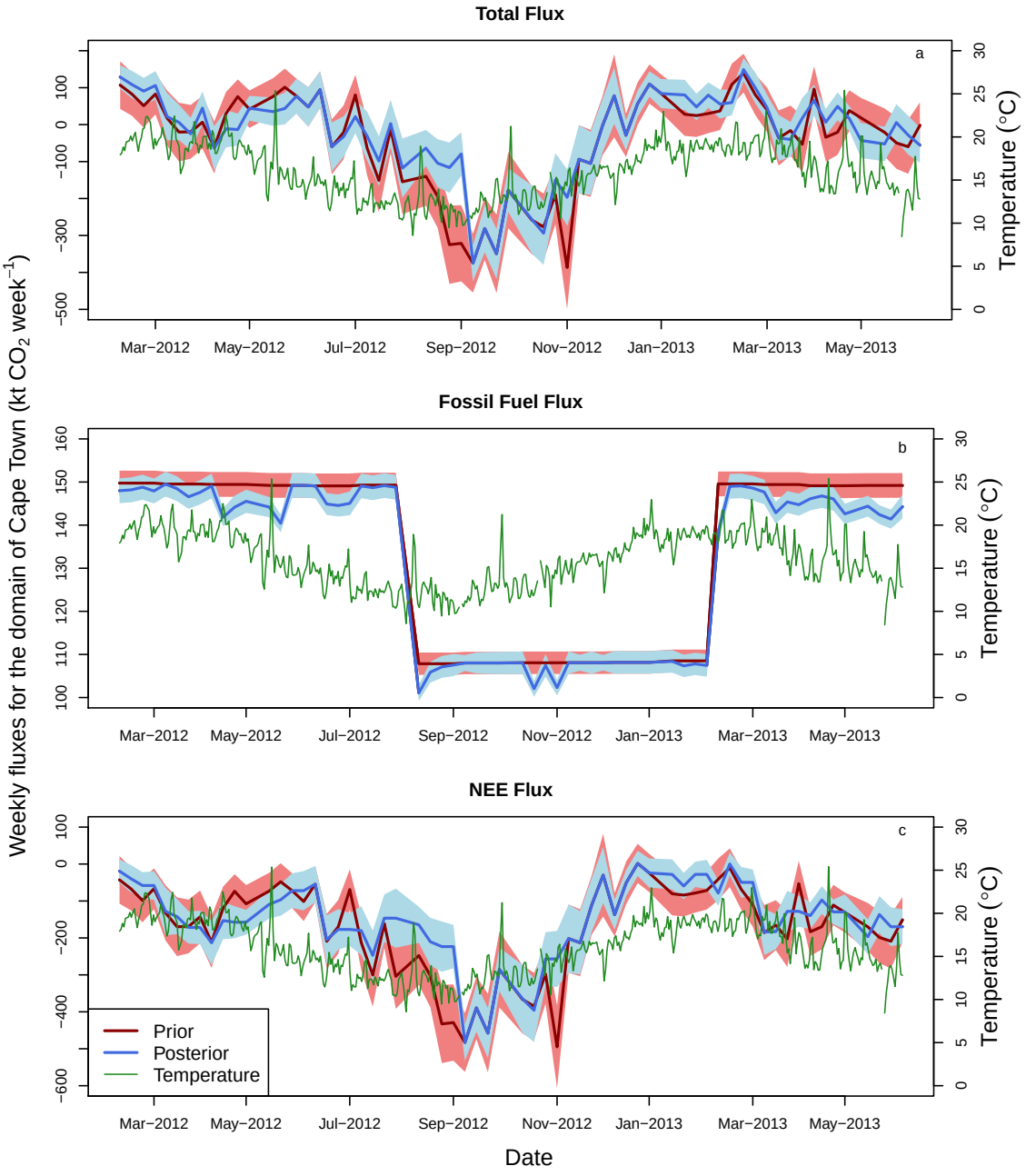


Figure 16. Prior (red) and posterior (blue) total weekly CO_2 flux estimates (kt CO_2 week^{-1}) (left-side axis) and uncertainty limits (shaded area), represented as a 95 % confidence interval, across the full domain of Cape Town, shown as (a) the total flux, (b) the fossil fuel flux, and (c) the NEE flux. The daily temperature (°C) as recorded at the Cape Point GAW station is provided by the lower green line (right-side axis).

4 Discussion and Conclusions

4.1 Inversion Innovation

The inversion was able to substantially improve the agreement between the prior and posterior modelled concentrations, with posterior modelled concentrations tracking most of the local events observed in the measurements. The most notable corrections to the pixel-level fluxes by the inversion were made to those with the largest industrial point sources, to pixels located on Robben Island where activities unaccounted for in the inventory were taking place, and to the areas dominated by NEE fluxes and located relatively close to the measurement sites. This evidence suggests that the inversion framework used here has had some success in capturing information regarding the CO_2 fluxes in the CT domain, and has applied reasonable corrections to the sources considered.

The inversion was able to reduce uncertainty of the total flux within a pixel by up to 97.7%, and was able to reduce the uncertainty in the total weekly flux over the whole domain by up to 50.5%. The largest innovation to a fossil fuel flux was applied to a pixel which contained an important point source in the domain - a crude oil refinery. This facility can process up to 100,000 barrels of crude oil per day. Unlike most industrial sources in the area which would be expected to have fairly consistent activity from day to day, activity at the crude oil refinery would depend on oil supply and on the global oil prices. During the period of March 2012 to June 2013, the global monthly oil price deviated between $117.29 in March 2012 and $90.73 in June 2012, ending on $99.74 in June 2013 (World Bank commodity prices). In addition, the consumption of liquid petroleum gas and heavy furnace oils may have decreased during this period (City of Cape Town, 2015). As this is a source with a large amount of expected variability, it is unlikely that the inversion method with distant measurement sites will be able to adequately estimate the flux in this pixel. In order for this to occur, the measurement site would have to consistently view the source during periods of both high and low activity in order to provide an unbiased estimate. An inversion could be used to estimate this particular source if a ring of instruments were placed around the site in order to capture information from the site at all times, regardless of the prevailing wind direction, such as suggested for the Otway CO_2CRC carbon capture project (Cook, 2012).

Compared with the fossil fuel emissions, relative innovations to the NEE fluxes were much larger, due to the large uncertainty prescribed to these fluxes. The largest innovations were made to natural areas near the CBD of CT, as well as to agricultural regions within the domain, particularly those close to the measurement sites. The prior estimates are dependent on the CABLE land-atmosphere interaction model and, although driven by the CCAM regional climate model, which in turn was driven by reliable reanalysed observations of the climate from NCEP, is still under ongoing development for use over South Africa. There is a great deal of uncertainty in its ability to simulate fluxes over the fynbos biome, as there is for most dynamic vegetation models (Moncrieff et al., 2015). In general, the inversion tended to increase the NEE fluxes so that the total flux was less negative compared with the priors, indicating that the amount of productivity estimated by CABLE may be overestimated.

4.2 Distinguishing Fossil Fuel and Biogenic Emissions

The spatial distribution and magnitude of the posterior fluxes and their uncertainties is strongly dependent on the prior spatial assignment of sources. In a city like CT, fossil fuel and NEE fluxes are usually co-occurring in the same pixel, with vegetation within the city representing a significant and large sink of CO_2. Under the current framework, if no fossil fuel source is prescribed to occur in a particular pixel, the inversion would only be able to adjust the NEE flux in this pixel, as the fossil fuel flux and its uncertainty are set to zero. If there is an unknown fossil fuel source in a pixel the inversion would lead to a better match between the modelled and observed concentrations, but a worse NEE flux estimate. Therefore the success of the inversion is largely dependent on how well the spatial extent of fossil fuel and biogenic sources are prescribed in the prior information.

In Bréon et al. (2015) NEE fluxes were aggregated to a larger grid size than the fossil fuel emissions. Effectively this means that perfect correlation was applied between the NEE fluxes for all pixels which fall within the same larger NEE pixel. By distinguishing the biogenic and fossil fuel sources in this way, it may allow the inversion to correctly allocate corrections between the fossil fuel and biogenic sources. We attempted to implement a similar idea by allowing correlation between NEE fluxes of neighbouring pixels and not prescribing correlations between fossil fuel sources. As the model tended to reduce fossil fuel emissions and increase NEE fluxes in the same pixel, it appears that the inversion is unlikely to adequately adjust the individual fluxes making up the total flux from a pixel. With the large coverage of vegetation within the domain, it is unlikely that a measurement network with only two sites could accurately estimate a given industrial point source, but there is still potential to monitor the overall city emissions, and assess the feasibility of inventory information. This is evident from the large uncertainty reductions attained by the inversion for the total weekly flux estimates aggregated over the domain.

The posterior uncertainty of any linear combination of terms from the control vector of the fluxes (including the difference between two fluxes from the same pixel) will always be reduced or (at worst) left unchanged relative to the prior uncertainty of the same linear combination of elements (Jackson, 1979; Jackson and Matsu'ura, 1985). This means that although negative correlation between the flux components may be introduced through the inversion as observations are made of the sum of these components, the uncertainty in both the difference between fluxes from the same pixel and the total flux within a pixel will be reduced.

If we define the distinction between the fossil fuel flux and NEE flux within the same pixel *i* as the variance of the difference between the fossil fuel and NEE fluxes $s_{f,i} - s_{NEE,i}$, this will be equal to the sum of the variances of these two fluxes minus twice the covariance between them: $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} - 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ will be negative. Therefore although the posterior uncertainty of the difference in these fluxes will always be larger than the sum of the individual posterior flux uncertainties, it will be smaller than the prior uncertainty of this linear combination of terms. The ability of the inversion under the current framework to distinguish between NEE and fossil fuel fluxes is limited as the posterior uncertainties are still large, and therefore the uncertainty of $s_{f,i} - s_{NEE,i}$ is large.

On the other hand, when we aggregate these fluxes from the same pixel to get the total flux within a pixel $s_{f,i} + s_{NEE,i}$, the uncertainty of this term is equal to $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} + 2 \times C_{s(f,i;NEE,i)}$ where $C_{s(f,i;NEE,i)}$ is negative. When we

aggregate fluxes from the same pixel, the uncertainty of this total is smaller due to the both the smaller posterior uncertainties of the individual fluxes and also because the covariances are negative. This demonstrates that the value of the inversion is to reduce the uncertainty on each of the individual fluxes and to additionally reduce the uncertainty of the aggregation of the NEE and fossil fuel flux within the same pixel. In this case, the reduction in the uncertainty of the sum of fluxes within the same pixel is strongly dependent on the size of the uncertainty of the NEE flux, which is usually the larger uncertainty.

We have shown in the supplementary material (section 1.8) that if we reduce the uncertainty in the modelled concentrations, the negative off-diagonal covariance terms of $\mathbf{C_s}$ become larger in magnitude. To improve the ability of the inversion to estimate the total flux within a pixel, we need to improve the skill of the atmospheric transport model, and we need to reduce the uncertainty in the estimates of the NEE. As it stands, with a large prior uncertainty in the estimation of the NEE fluxes from the CABLE model which remains a large posterior estimate after the inversion, the distinction between the fossil fuel and NEE flux from the same pixel, $C_{s(f,i;f,i)} + C_{s(NEE,i;NEE,i)} - 2 \times C_{s(f,i;NEE,i)}$, is not very different from the prior estimate, $C_{s_0(f,i;f,i)} + C_{s_0(NEE,i;NEE,i)}$.

4.3 Strengths and Limitations

This paper represents a first attempt at estimating $\mathrm{CO_2}$ fluxes at the high resolution of $1\,\mathrm{km}$ by $1\,\mathrm{km}$ over CT, solving for individual fossil fuel, terrestrial and oceanic biogenic sources. A previous network design study for South Africa (Nickless et al., 2015b), which aggregated the NEE and fossil fuel fluxes up to a $15\,\mathrm{km} \times 15\,\mathrm{km}$ resolution, showed the aggregation errors could be high at the regional level. By avoiding this aggregation, and maintaining the $1\,\mathrm{km}$ by $1\,\mathrm{km}$ resolution of the atmospheric transport model throughout the inversion process, we attempted to minimise these aggregation errors. Maintaining this resolution is computationally expensive, but possible due to the relatively small domain size of $100\,\mathrm{km}$ by $100\,\mathrm{km}$.

A limitation of this study is that human respiration was not explicitly accounted for. With a population of over three million, this flux could represent up to $26\,\mathrm{kt\ CO_2\ week^{-1}}$, if we attribute $1\,\mathrm{kilogram\ CO_2}$ per day to each person (Bréon et al., 2015). This represents between a fifth to a quarter of the total fossil fuel flux estimated for the domain, and therefore is by no means a negligible quantity. Including this fossil fuel flux in the inventory information would most likely lead to confounding between the domestic emissions and the human respiration, as these two sources would have been calculated based on population data. Domestic emissions were heavily dependent on the assumption regarding how domestic heating emissions were distributed during the year. This had a large impact on the temporal profile of fossil fuel emissions, resulting in a lower average emission in summer compared with winter, which persisted in the posterior estimates of the fossil fuel fluxes. Consequently the prior uncertainties in the domestic emissions were larger compared with other fossil fuel sources. We have included sensitivity analyses on the specification of the fossil fuel emissions and their uncertainties in a companion paper.

This inversion was performed by solving for the total observed concentrations, rather than solving for the gradient in concentration measurements between two sites (Bréon et al., 2015). There are several reasons why the gradient method would not have been suitable in these circumstances. Firstly, we had the advantage of a background site which viewed background levels over 70% of the time. As our city is located within the Southern Hemisphere, far less variability is expected between the hourly measurements or from week to week compared with the Northern Hemisphere, and we would expect the $\mathrm{CO_2}$

concentration at the oceanic boundaries of our domain to be similar to background levels of CO_2. This was confirmed by the results of the inversion, which made almost no adjustments to the oceanic boundary concentrations, but made slightly larger adjustments to the boundary concentrations at the north and east terrestrial boundaries. This implies that the adjustments made by the inversion were largely due to the surface flux sources within the domain. In addition, there are no large expected sources located anywhere near the boundaries of the domain. The next major city in the proximity of CT is Port Elizabeth, which is over 600 km away. For this reason, it is unnecessary to solve for the gradients between the two sites, as we did not require the removal of unknown outside sources from the observation dataset. This allowed us to use the entire measurement record, which is an advantage as we only had two CO_2 measurement sites available.

Secondly, the gradient method would likely have performed poorly here, as the direction of travel of an air parcel between the two sites would not necessarily be in a straight line due to the topography of the site and demonstrated by the plot of the sensitivities at the two sites to the domain (Figure 17), as well as the modelled wind fields provided in supplementary material section 1.3. Therefore extracting observations based on the prevailing wind speed and directions at the sites would have not represented true gradients in the CO_2 concentrations between the two measurements sites.

These sensitivities of the sites to the surface fluxes also reveal that the sites are often viewing oceanic sources (Figure 17). A limitation of this study is that a single time series of ocean biogenic emissions was used as the prior estimate for all oceanic pixels in the domain. The fluxes from the near-coastal oceanic pixels are likely to have significant spatial heterogeneity, although smaller compared with the terrestrial biogenic fluxes. A way of improving this would be to use the output of a model representing atmosphere-land-ocean biogeochemical exchanges to provide prior fluxes over both the land and ocean. The CSIR's Variable Resolution Earth System Model (VRESM) is such a model currently under development, which aims to couple CCAM, CABLE, and CSIR's Variable-Cubic Ocean Model (VCOM), and Pelagic Interactions Scheme for Carbon and Ecosystem Studies (PISCES) to model ocean transport and biogeochemistry (Engelbrecht et al., 2016). Due to the amount of shipping activity around the CT harbour and within the ocean domain viewed by the Robben Island site, the inventory analysis could be improved by the inclusion of information on shipping routes so that the inversion can adjust fossil fuel fluxes in these ocean pixels as well.

The uncertainty in the NEE fluxes played an important role in the outcome of the inversion. If tighter uncertainty limits could be placed on the NEE flux estimates from the land-atmosphere exchange model, it would allow the inversion to better distinguish between the fossil fuel and NEE fluxes. This could be attained by validation work through eddy-covariance flux measurements over dominant vegetation types within the domain, for example. The use of a land-atmosphere exchange model well suited to the vegetation within the domain of a city-scale inversion is essential for improving the ability of the inversion to adjust the fossil fuel sources which are of foremost interest. The dependence on knowing the NEE fluxes well in order to estimate the fossil fuel fluxes could be reduced if there were additional measurements of Δ^{14}C and δ^{13}C isotope measurements at each of the sites, including at the background site (Turnbull et al., 2015; Newman et al., 2016).

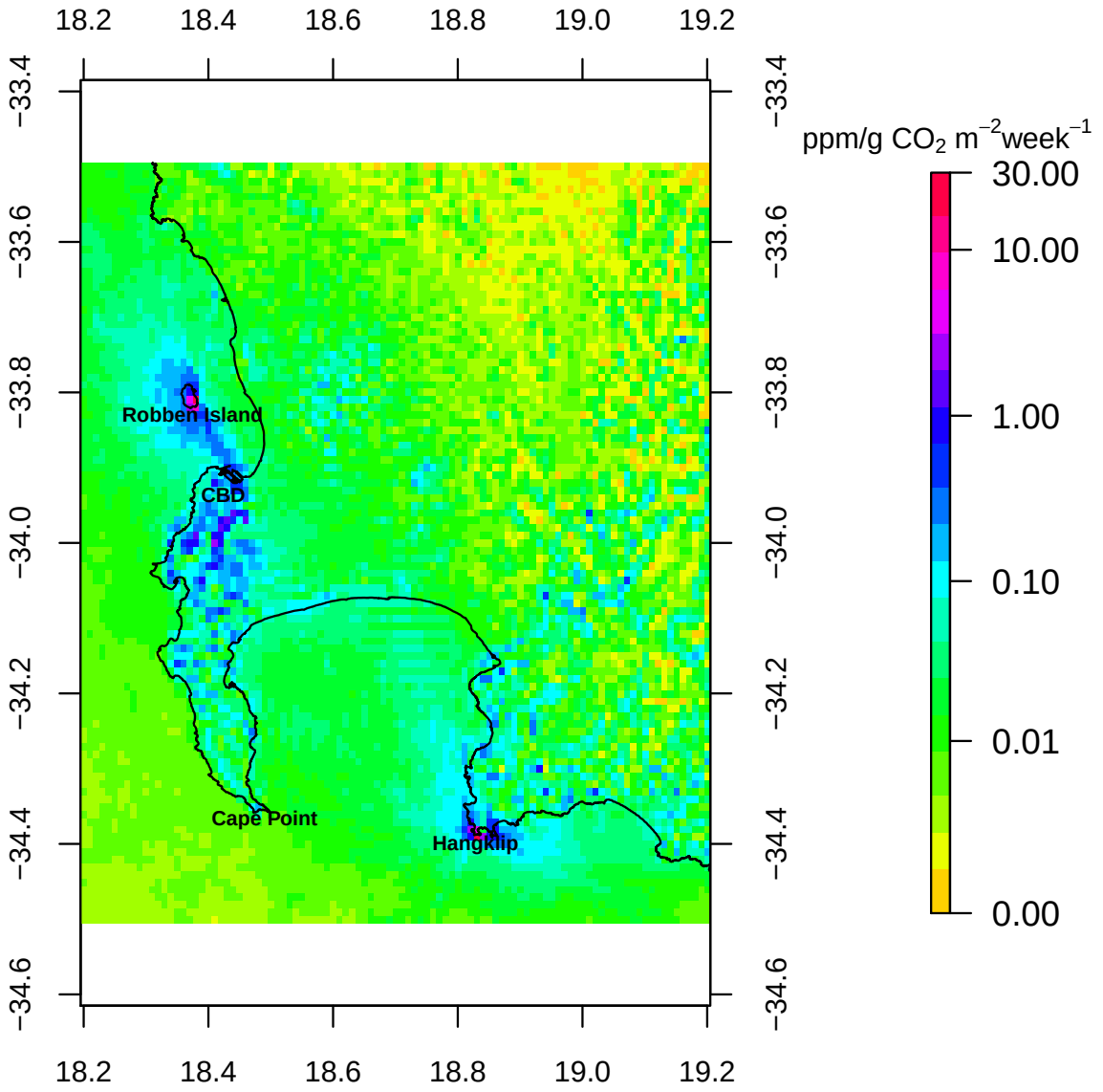


Figure 17. Mean weekly sensitivities ($\mathrm{ppm\,kg^{-1}\ CO_2 m^{-2}\,week^{-1}}$) of the measurement sites at Robben Island and Hangklip to each surface source pixel, plotted on a log scale.

5 Conclusions

We have presented the results of a city-scale atmospheric inversion for Cape Town, South Africa. We have shown that the current inversion framework was able to reduce the misfit between the observed and modelled concentrations by making reasonable adjustments to the fossil fuel and NEE fluxes. We were able to reduce the total uncertainty in CO_2 weekly flux from the domain by up to 50%. We have demonstrated the ability of an inversion to obtain an improved estimate of an aggregated flux, even when the uncertainty in one component is large. To realistically use this framework for an operational inversion system for use in MRV, the uncertainty around the NEE estimates for Cape Town needs to be reduced. Further qualifying concentration observations according to the contributions from anthropogenic sources and those from the biosphere will help the inversion to disentangle the corrections to these two fluxes, and will reduce the dependency of the fossil fuel flux corrections on the uncertainty of the NEE flux estimates.

Author contributions. AN installed and maintained all the instrumentation at Robben Island and Hangklip, obtained the measurements and processed these into hourly concentrations, ran and processed the result of the LPDM in Fortran, produced all code and ran the inversion in Python, processed all the inversion results using R Statistical Software, produced all graphics and tables, and was responsible for the development of the manuscript which forms part of her PhD. PJR was the main scientific supervisor, oversaw all implementation of the inversion, and provided guidance on the presentation of results. E-GB provided the Cape Point concentration and meterological data, and provided assistance and guidance in the installation and maintenance of the Picarro instrumentation. FE provided the CCAM and CABLE data. BE provided guidance on statistical issues. RJS provided guidance on the location of the sites and provided advice on the interpretation of the biogenic fluxes. All authors had the opportunity to comment on the manuscript.

Competing interests. The authors declare that they have no conflict of interest

Code and data availability. Data and code related to the Bayesian inversion procedure can be made available upon request

Acknowledgements. We would like to acknowledge and thank Dr. Casper Labuschagne and Danie van der Spuy of the South African Weather Service for their assistance in maintaining the instruments at Robben Island and Hangklip, and Dr. Casper Labuschagne for his guidance on processing the instantaneous CO_2 concentration data; Martin Steinbacher for providing guidance and schematics on the calibration system used on the Picarro instruments; Robin Poggenpoel and Jacobus Smith of Transnet for allowing us access to the lighthouses; Peter Saaise of Transnet, the Robben Island lighthouse keeper, (and his daughter) for assisting when the instrument was not responding; Marek Uliasz for providing us access to the code for his LPDM model; Thomas Lauvaux for providing guidance on processing the LPDM results and useful discussion on the boundary contribution in the inversion; Keith Moir, Rooi Els, for providing wind data near Hangklip. Use was made of the University of Cape Town ICTS-HPC cluster. Please see http://hpc.uct.ac.za/ for details. We would like to thank Andrew Lewis of

the University of Cape Town HPC facility for providing useful advice on improving the efficiency of the Python runs. This research was funded by competitive parliamentary grant funding from the Council of Scientific and Industrial Research awarded to the Global Change Competency Area towards the development of the Variable-resolution Earth System Model (VRESM; Grants EEGC030 and EECM066). Additional funding was obtained from the South African National Research Foundation for the Picarro instrumentation.

References

Andres, R. J., Boden, T. A., Bréon, F. -M., Ciais, P., Davis, S., Erickson, D., Gregg, J. S., Jacobson, A., Marland, G., Miller, J., Oda, T., Olivier, J. G. J., Raupach, M. R., Rayner, P., and Treanton, K.: A synthesis of carbon dioxide emissions from fossil-fuel combustion, Biogeosciences, 9, 1845–1871, doi: 10.5194/bg-9-1845-2012, 2012.

Andres, R. J., Boden, T. A., and Higdon, D.: A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission, Tellus B, 66, 23616, doi: 10.3402/tellusb.v66.23616, 2014.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bourquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., Zhu, Z.: TransCom 3 inversion intercomparison: impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi: 10.1029/2004GB002439, 2006.

Bellassen, V. and Stephan, N.: Accounting for carbon: Monitoring, reporting and verifying emissions in the climate economy, Cambridge University Press, Cambridge, UK, 2015.

Boden, T. A., Marland, G., and Andres, R. J.: Global, regional, and national fossil fuel CO_2 emissions, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A., doi: 10.3334/CDIAC/00001_V2011, 2011.

Boon, A., Broquet, G., Clifford, D. J., Chevallier, F., Butterfield, D. M., Pison, I., Ramonet, M., Paris, J. -D., and Ciais, P.: Analysis of the potential of near-ground measurements of CO_2 and CH_4 in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model, Atmos. Chem. Phys., 16, 6735–6756, doi:10.5194/acp-16-6735-2016, 2016.

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual atmospheric CO_2 sources and sinks: 1. Method and control inversion, J. Geophys. Res., 104, 26161–26178, 1999.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707–1724, doi: 10.5194/acp-15-1707-2015, 2015.

Brioude, J., Angevine, W. M., Ahmadov, R., Kim, S. -W., Evan, S., McKeen, S. A., Hsie, E. -Y., Frost, G. J., Neuman, J. A., Pollack, I. B., Peischl, J., Ryerson, T. B., Holloway, J., Brown, S. S., Nowak, J. B., Roberts, J. M., Wofsy, S. C., Santoni, G. W., Oda, T., and Trainer, M.: Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NOxand CO_2 and their impacts, Atmos. Chem. Phys., 13, 3661–3677, doi: 10.5194/acp-13-3661-2013, 2013.

Brunke, E. -G, Labuschagne, C., Parker, B., Scheel, H. E, and Wittlestone, S.: Baseline air mass selection at Cape Point,South Africa: application of ^{222}Rn and other filter criteria to CO_2, Atmos. Environ., 38, 5693–5702, doi:10.1016/j.atmosenv.2004.04.024, 2004.

Carslaw, D. C. and Ropkins, K.: openair — an R package for air quality data analysis, Environ. Model. Softw., 27–28, 52–61, doi:10.1016/j.envsoft.2011.09.008, 2012.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morgui, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res., 115, D21307, doi: 10.1029/2010JD013887, 2010.

Chouinard, C., Béland, M., McFarlane, N.: A simple gravity wave drag parameterization for use in medium-range weather forecast models, Atmos. Ocean, 24, 91–110, 1986.

Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., and Ramonet, M.: Atmospheric inversions for estimating CO_2 fluxes: methods and perspectives, Climatic Change, 103, 69–92, 2010.

City of Cape Town: State of energy and energy futures report. Cape Town: City of Cape Town, http://www.capetown.gov.za/en/EnvironmentalResourceManagement/publications/Documents/State_of_Energy_+_Energy_Futures_Report_2011_revised_2012-01.pdf, 2011.

City of Cape Town: State of energy report 2015. Cape Town: City of Cape Town, https://africancityenergy.org/getfile.php?id=22&category=0, 2015.

P. Cook (Ed.): Geologically storing carbon: The CO2CRC Otway Project: Learning from the Otway Project experience, CSIRO Press, Melbourne, p. 384, 2014.

Crosson, E.:A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, Appl. Phys. B, 92, 403–408, doi:10.1007/s00340-008-3135-y, 2008.

Davis, K. J., Deng, A., Lauvaux, T., Miles, N. L., Richardson, S. J., Sarmiento, D. P., Gurney, K. R., Hardesty, R. M., Bonin, T. A., Brewer, W. A., Lamb, B. K., Shepson, P. B., Harvey, R. M., Cambaliza, M. O., Sweeney, C., Turnbull, J. C., Whetstone, J., and Karion, A.: The Indianapolis Flux Experiment (INFLUX): A test-bed for developing urban greenhouse gas emission measurements, Elem. Sci. Anth., 5:21, doi: 10.1525/elementa.188, 2017.

UK Department for Environment, Food and Rural Affairs (Defra): UK ship emissions inventory. Final report. London: Crown, http://uk-air.defra.gov.uk/assets/documents/reports/cat15/1012131459_21897_Final_Report_291110.pdf, 2010.

UK Department for Environment, Food and Rural Affairs (Defra): Government GHG conversion factors for company reporting: Methodology paper for emission factors, London: Crown, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224437/pb13988-emission-factor-methodology-130719.pdf, 2013.

UK Department for Environment, Food and Rural Affairs (Defra): Treatment of uncertainties for national estimates of greenhouse gas emissions, http://uk-air.defra.gov.uk/reports/empire/naei/ipcc/uncertainty/contents.html, 2013.

Dlugokencky, E. and Tans, P.: Trends in atmospheric carbon dioxide, National Oceanic and Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL), available at: http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html, last access: 29 September 2016, 2016.

Duren, R. M. and Miller, C. E.: Measuring the carbon emissions of megacities, Nat. Clim. Change, 2, 560–562, doi: 10.1038/nclimate1629, 2012.

Engelbrecht, F. A., McGregor, J. L. and Engelbrecht, C. J.: Dynamics of the conformal-cubic atmospheric model projected climate-change signal over southern Africa, Int. J. Climatol., 29, 1013-1033., doi: 10/1002/joc.1742.29., 2009.

Engelbrecht, F. A., Landman, W. A., Engelbrecht, C. J., Landman, S., Bopane, M. M., Roux, B., McGregor, J. L., and Thatcher, M.: Multi-scale climate modelling over Southern Africa using variable-resolution global model, Water Research Commission 40-Year Celebration Conference, Kempton Park, 31 August - 1 September 2011, doi: 10.4314/wsa.v37i5.2, 2011.

Engelbrecht, C. J., Engelbrecht, F. A. and Dyson, L. L.: High-resolution model-projected changes in mid-tropospheric closed-lows and extreme rainfall events over southern Africa. Int. J. Climatol., 33, 173–187, doi: 10/1002/joc.3420, 2013.

Engelbrecht, F., Adegoke, J., Bopape, M-J., Naidoo, M., Garland, R., Thatcher, M., McGregor, J., Katzfey, J., Werner, M., Ichoku, C. and Gatebe, C.: Projections of rapidly rising surface temperatures over Africa. Env. Res. Letters., 10(8), 085004, doi: 10.1088/1748-9326/10/8/085004, 2015.

Engelbrecht, F. A., McGregor, J. L., Thatcher, M., Garland, R., Sovara, M., Bopane, M. M., and van der Merwe, J.: The Variable-Resolution Earth System Model and its simulations of the Benguela upwelling system. The International Conference on Regional Climate CORDEX 2016, Stockholm, Sweden, 17–20 May 2016, 2016.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge Univ. Press, New York, 2002.

Erickson, P. and Tempest, K.: Advancing climate ambition: Cities as partners in global climate action, Produced by Stockholm Environment Institute (SEI) in support of the UN Secretary-General's Special Envoy for Cities and Climate Change and C40, 2014.

Exbrayat, J. -F., Pitman, A. J. Abramowitz, G. and Wang, Y, -P.: Sensitivity of net ecosystem exchange and heterotrophic respiration to parameterization uncertainty, J. Geophys. Res. Atmos., 118, 1640–1651, doi:10.1029/2012JD018122, 2013.

Fawcett, R. A., Pitcher, G., Bernard, S., Cembella, A., and Kudela, R.: Contrasting wind patterns and toxigenic phytoplankton in the southern Benguela upwelling system, Mar. Ecol. Prog. Ser., 348, 19–31, doi: 10.3354/meps07027, 2007.

Feng, S., Lauvaux T., Newman, S., Rao, P., Ahmadov, R. Deng, A., Díaz-Isaac, L. I., Duren, R. M., Fischer, M. L., Gerbig, C., Gurney, K. R., Huang, J., Jeong, S., Li, Z., Miller, C. E., O'Keeffe, D., Patarasuk, R., Ser, S. P., Song, Y., Wong, K. W., and Yung, Y. L.: Los Angeles megacity: a high-resolution land–atmosphere modelling system for urban CO_2 emissions, Atmos. Chem. Phys., 16, 9019–9045, doi: 10.5194/acp-16-9019-2016, 2016.

Ganesan, A. L., Rigby, M., Zammit-Mangion, A., Manning, A. J., Prinn, R. G., Fraser, P. J., Harth, C. M., Kim, K. -R., Krummel, P. B., Li, S., Mühle, J., O'Doherty, S. J., Park, S., Salameh, P. K., Steele, L. P., and Weiss, R. F.: Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, Atmos. Chem. Phys., 14, 3855-3864, doi:10.5194/acp-14-3855-2014, 2014.

Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rödenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe – Part 1: mapping the atmospheric CO_2 signals, Atmos. Chem. Phys., 7, 3461-3479, doi: 10.5194/acp-7-3461-2007, 2007.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S. and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, J. Geophys. Res., 108(D24), 4756, doi: 10.1029/2002JD003018, 2003.

Gregor, L. and Monteiro P. M. S.: Is the southern Benguela a significant regional sink of CO2? S. Afr. J. Sci., 109(5/6), Art. $\#0094$, 5 pages, doi: 10.1590/sajs.2013/20120094, 2013.

Gurney. K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 405, 626–630, 2002.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Kowalczyk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.: TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, Tellus B, 55, 555–579, 2003.

Gurney, K. R., Mendoza, D. L., Zhou, Y. Y., Fischer, M. L., Miller, C. C., Geethakumar, S., Du Can, S. D.: High resolution fossil fuel combustion CO_2 emission fluxes for the United States. Environ. Sci. Technol., 43 (14), 5535–5541, doi: 10.1021/es900806c, 2009.

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., and Abdul–Massih, M.: Quantification of fossil fuel CO_2 emissions on the building/street scale for a large U.S. city, Environ. Sci. Technol., 46, 12194–12202, doi: 10.1021/es3011282, 2012.

Holtslag, A. A. M. and Boville, B. A.: Local versus non-local boundary layer diffusion in a global climate model, J. Clim., 6, 1825–1842, doi: 10.1175/1520-0442(1993)006<1825:LVNBLD>2.0.CO;2, 1993.

Hutyra, L., Duren, R., Gurney, K. R., Grimm, N., Kort, E. A., Larson, E., Shrestha, G.: Urbanization and the carbon cycle: Current capabilities and research outlook from the natural sciences perspective. Earth's Future, 2: 473–495. doi: 10.1002/2014EF000255, 2014.

Intergovernmental Panel on Climate Change (IPCC): Good practice guidance and uncertainty management in national greenhouse gas inventories. Montreal: IPCC, 93–102, http://www.ipcc-nggip.iges.or.jp/public/gp/english/, 2000.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K. and Meyer, L.A. (Eds.)]. IPCC, Geneva, Switzerland, 151 pp, 2014.

Jackson, D. D.: The use of a priori data to resolve non-uniqueness in linear inversion, Geophys. J. R. astr. Soc., 57, 137–157, 1979.

Jackson, D. D and Matsu'ura, M.: A Bayesian approach to nonlinear inversion, J. Geophys. Res. 90 (B1), 581–591, 1985.

Janssens-Maenhout, G., Pagliari, V., Guizzardi, D., and Muntean, M.: Global emission inventories in the Emission Database for Global Atmospheric Research (EDGAR) – Manual (I). Gridding: EDGAR emissions distribution on global gridmaps. Joint Research Centre, Luxembourg: European Union, 33 pp, doi: 10.2788/81454, 2012.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K., C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds R., Jenne, R., Joseph, D.: The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteorol. Soc., 77, 437–472, 1996.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three dimensional global inverse model of the atmospheric transport, 2. Inversion of the transport of CO_2 in the 1980s, J. Geophys. Res., 104, 18555–18581, 1999.

Kaminski, T., Rayner, P J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4705–4715, 2001.

Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO_2 emissions: minimum site location requirements for the Los Angeles megacity, J. Geophys. Res., 118, 1–8, doi: 10.1002/jgrd.50135, 2013.

Kowalczyk, E. A., Garratt, J. R. and Krummel, P. B.: Implementation of a soil-canopy scheme into the CSIRO GCM - regional aspects of the model response, CSIRO Div. Atmospheric Research, Melbourne, Australia, Tech Paper No. 32, 59 pp., 1994.

Kowalczyk, E. A., Wang, Y. P. and Law, R. M.: CSIRO Atmosphere Biosphere Land Exchange model for use in climate models and as an offline model, CSIRO Marine and Atmospheric Research technical paper xxV ISBN 1 921232 39 0, 2006.

Lacis, A. A. and Hansen, J. E.: A parameterization for the absorption of solar radiation in the earth's atmosphere, J. Atmos. Sci., 31, 118–133, doi: 10.1175/1520-0469(1974)031<0118:APFTAO>2.0.CO;2, 1974.

Landman, S., Engelbrecht, F. A, Dyson, L., Engelbrecht, C., J., and Landman, W., A.: A short-range ensemble prediction system for South Africa, Water SA, 38(5), 765–774, doi: 10.4314/wsa.v38i5.16, 2012.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. Phys., 8, 3459–3471, doi: 10.5194/acp-8-3459-2008, 2008.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, doi: 10.5194/bg-6-1089-2009, 2009.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337–354, doi: 10.5194/acp-12-337-2012, 2012.

Lauvaux, T., Miles, N. L., Richardson, S. J., Deng, A., Stauffer, D. R., Davis, K. J., Jacobson, G., Rella, C., Calonder, G. -P., and DeCola, P. L.: Urban emissions of CO_2 from Davos, Switzerland: the first real-time monitoring system using atmospheric inversion technique, J. Appl. Meteorol. Climatol., 52, 2654–2668, doi: 10.1175/JAMC-D-13-038.1, 2013.

Lauvaux, T., Miles, N. L., Deng, A. Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), J. Geophys. Res. Atmos., 121, 5213–5236, doi: 10.1002/2015JD024473, 2016.

Law, R. M., Chen, Y., Gurney, K. R., and Transcom 3 Modellers: TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, Tellus B, 55, 580–595, 2003.

Law, R. M., Raupach, M. R., Abramowitz, G., Dharssi, I., Haverd, V.,Pitman, A. J.,Renzullo, L., Van Dijk, A. and Wang, Y. -P.: The Community Atmosphere Biosphere Land Exchange (CABLE) model Roadmap for 2012-2017, CAWCR Technical Report No. 057, 2012.

Le Quéré, C., Andres, R. J., Boden, T. A., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D.,Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, Earth Syst. Sci. Data, 5, 165–185, doi: 10.5194/essd-5-165-2013 2013.

McGregor, J. L., Gordon, H. B., Watterson, I. G., Dix, M. R., and Rotstayn, L. D.: The CSIRO 9-level atmospheric general circulation model, CSIRO Div. Atmospheric Research Tech, Paper No. 26, 89 pp., 1993.

McGregor, J. L.: Semi-Lagrangian advection on conformal-cubic grids. Mon. Weather Rev., 124, 1311–1322, doi: 10.1175/1520-0493(1996)124<1311:SLAOCC>2.0.CO;2, 1996.

McGregor, J. L. and Dix, M. R.: The CSIRO conformal-cubic atmospheric GCM, in: IUTAM Symposium on Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics, Limerick, Ireland, 2–7 July 2000, edited by: Hodnett, P. F., Kluwer, Dordrecht, 197–202, 2001.

McGregor, J. L.: A new convection scheme using a simple closure, in: Current issues in the parameterization of convection, BMRC, Melbourne, Australia, Research Report 93, 33-36, 2003.

McGregor, J. L.: Geostrophic adjustment for reversibly staggered grids. Mon. Weather Rev., 133, 1119–1128, doi: 10.1175/MWR2908.1, 2005a.

McGregor, J. L.: C-CAM: Geometric aspects and dynamical formulation. CSIRO Atmospheric Research Technical Paper, No 70, 41, 2005b.

McGregor, J. L. and Dix, M. R.: An updated description of the Conformal-Cubic Atmospheric Model, in: High Resolution Numerical Modelling of the Atmosphere and Ocean, edited by: Hamilton, K. and Ohfuchi, W., Springer, New York, USA, 51–76, 2008.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, Proc. Natl. Acad. Sci. U. S. A., 109(22), 8423–8428, doi: 10.1073/pnas.1116645109, 2012.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 110, D24107, doi: 10.1029/2005JD005970, 2005.

Moncrieff, G.R., Scheiter,S. Slingsby, J. A. and Higgins, S. I.: Understanding global change impacts on South African biomes using Dynamic Vegetation Models, S. Afr. J. Bot., 101, 16–23, doi: 10.1016/j.sajb.2015.02.004, 2015.

Newman, S., Xu, X., Gurney, K. R., Hsu, Y. K., Li, K. F., Jiang, X., Keeling, R., Feng, S., O'Keefe, D., Patarasuk, R., Wong, K. W., Rao, P., Fischer, M. L., and Yung, Y. L.: Toward consistency between trends in bottom-up CO_2 emissions and top-down atmospheric measurements in the Los Angeles megacity, Atmos. Chem. Phys., 16, 3843–3863, doi: 10.5194/acp-16-3843-2016, 2016.

Nickless, A., Scholes, R. J. and Filby, E.: Spatial and temporal disaggregation of anthropogenic CO_2 emissions from the City of Cape Town, S. Afr. J. Sci., 111(11/12), Art. $\#2014-0387$, 8 pages, doi: 10.17159/sajs.2015/20140387, 2015.

Nickless, A., Ziehn, T., Rayner, P. J., Scholes, R .J., and Engelbrecht, F.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 2: Sensitivity analyses and South African test case, Atmos. Chem. Phys., 15, 2051–2069, doi: 10.5194/acp-15-2051-2015, 2015.

NRC (Committee on Methods for Estimating Greenhouse Gas Emissions): Verifying greenhouse gas emissions: methods to support international climate agreements (9780309152112), The National Academies Press, Washington, D.C., 2010.

Oda, T., Lauvaux, T., Lu, D., Rao, P., Miles, N. L., Richardson, S. J. and Gurney, K. R.: On the impact of granularity of space-based urban CO_2 emissions in urban atmospheric inversions: A case study for Indianapolis, IN, Elem Sci Anth, 5, 28, doi: 10.1525/elementa.146, 2017.

Peylin, P., Baker, D., Sarmiento, J., Ciais, P., and Bousquet, P.: Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO_2 data, J. Geophys. Res., 107, 4385, doi: 10.1029/2001JD000857, 2002.

Rayner, P. J., Enting, I. G., Francey, R. J., and Langenfelds, R. L.: Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, Tellus B, 51, 213–232, 1999.

Rayner, P. J., Law, R. M, Allison, C. E., Francey, R. J., Trudinger, C. M., Pickett-Heaps, C.: Interannual variability of the global carbon cycle (1992-2005) inferred by inversion of atmospheric CO_2 and $\delta^{13}C$ measurements, Global Biogeochem. Cy., 22, GB3008, doi: 10.1029/2007GB003068, 2008.

Rotstayn, L. D.: A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I: Description and evaluation of the microphysical processes, Q. J. R. Meteorol. Soc., 123, 1227–1282, 1997.

Roux, B.: Ultra high-resolution climate simulations over the Stellenbosch wine producing region using a variable-resolution model, MSc Thesis, Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa, 106 pp., 2009.

Schmidt, F.: Variable fine mesh in spectral global model, Beitr. Phys. Atmos., 50, 211–217, 1977.

Schuh, A. E., Lauvaux, T., West, T. O., Denning, A. S., Davis, K., J., Miles, N., Richardson, S., Uliasz, M., Lokupitiya, E., Cooley, D., Andrews, A., and Ogle, S.: Evaluating atmospheric CO_2 inversions at multiple scales over a highly inventoried agricultural landscape, Glob. Change Biol., 19, 1424–1439, doi: 10.1111/gcb.12141, 2013.

Schwarzkopf, M. D. and Fels, S. B.: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes, J. Geophys. Res., 96, 9075–9096, 1991.

Shiga, Y. P., Michalak, A. M., Gourdji, S. M., Mueller, K. L., and Yadav, V.: Detecting fossil fuel emissions patterns from subcontinental regions using North American in situ CO_2 measurements, Geophys. Res. Lett., 41(12), 4381–4388, doi: 10.1002/2014GL059684, 2014.

Shrout, P. E. and Fleiss, J. L.: Intraclass correlations: uses in assessing rater reliability, Psychol. Bull., 2, 420–428, 1979.

Seibert, P. and Frank, A.: Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode, Atmos. Chem. Phys., 4, 51–63, doi: 10.5194/acp-4-51-2004, 2004.

South African Department of Energy: Digest of South African energy statistics, Pretoria: Department of Energy. http://www.energy.gov.za/files/media/explained/2009%20Digest%20PDF%20version.pdf, 2009.

Statistics South Africa: Census 2011 statistical release, P0301.4., Pretoria: Statistics South Africa, 2011.

Staufer, J., Broquet, G., Bréon, F. -M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO2 emissions based on atmospheric inversion, Atmos. Chem. Phys., 16, 14703–14726, doi: 10.5194/acp-16-14703-2016, 2016.

Strong, C., Stwertka, C., Bowling, D. R., Stephens, B. B. and Ehleringer, J. R.: Urban carbon dioxide cycles within the Salt Lake Valley: A multiple-box model validated by observations, J. Geophys. Res., 116, D15307, doi: 10.1029/2011JD015693, 2011.

Seto, K. C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J. E., Müller, D. B., Murakami, J., Nagendra, H., and Ramaswami, A.: Human settlements, infrastructure and spatial planning, in: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlömer, S., von Stechow, C., Zwickel, T., and Minx, J. C., Cambridge, United Kingdom and New York, NY, USA, 2014.

Sugar, L. and Kennedy, C.: A low carbon infrastructure plan for Toronto, Canada, Can. J. Civ. Eng., 40, 86–96, doi: 10.1139/cjce-2011-0523, 2013.

Tans, P. and Keeling, R.: Mauna Loa CO_2 monthly mean data, NOAA/ESRL www.esrl.noaa.gov/gmd/ccgg/trends/ and Scripps Institution of Oceanography scrippsco2.ucsd.edu/, 2016.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, 2005.

Thatcher, M. and McGregor, J. L.: Using a scale-selective filter for dynamical downscaling with the conformal cubic atmospheric model, Mon. Weather Rev., 137, 1742–1752, 2009.

Thatcher, M. and McGregor, J. L.: A technique for dynamically downscaling daily-averaged GCM datasets over Australia using the Conformal Cubic Atmospheric Model, Mon. Weather Rev., 139, 79–95, 2010.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., and Patarasuk, P.: Toward quantification and source sector identification of fossil fuel CO_2 emissions from an urban area: Results from the INFLUX experiment, J. Geophys. Res. Atmos., 120, 292–312, doi: 10.1002/2014JD022555, 2015.

Uliasz, M.: The atmospheric mesoscale dispersion modeling system, J. Appl. Meteorol., 31, 139–149, 1993.

Uliasz, M.: Lagrangian particle modeling in mesoscale applications, in: Environmental Modelling II, Computational Mechanics Publications, Southampton, UK, 71–102, 1994.

UN–Habitat: Cities and climate change: global report on human settlements 2011, Earthscan, United Nations Human Settlements Programme (UN–Habitat), 2011.

Wang, Y, P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., van Gorsel, E. and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, J. Geophys. Res., 116, G01034, doi: 10.1029/2010JG001385, 2011.

Whittlestone, S., Kowalczyk, E., Brunke, E. G., and Labuschagne, C.: Source Regions for CO_2 at Cape Point Assessed by Modelling 222Rn and Meteorological Data, Technical Report for the South African Weather Service, Pretoria, South Africa, 2009.

Wu, L., Bocquet, M., Chevallier, F., Lauvaux, T., and Davis, K.: Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions, Tellus B, 65, 20894, doi: 10.3402/tellusb.v65i0.20894, 2013.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO_2 emissions?, Atmos. Chem. Phys., 16, 7743–7771, doi: 10.5194/acp-16-7743-2016, 2016.

Zhang, L., Zhang, H. and Li, Y.: Surface energy, water and carbon cycle in China simulated by the Australian community land surface model (CABLE), Theor. Appl. Climatol.,96(3),375–394, doi:10.1007/s00704-008-0047-z, 2009.

Zhang, H ., Pak, .B., Wang, Y. P., Zhou, X., Zhang, Y. and Zhang, L.: Evaluating surface water cycles simulated by the Australian community land surface model (CABLE) across different spatial and temporal domains. J. Hydrometeor, 14, 1119–1138, 2013.

Ziehn, T., Nickless, A., Rayner, P. J., Law, R. M., Roff, G., and Fraser, P.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 1: Methodology and Australian test case, Atmos. Phys. Chem., 14, 9363–9378, doi: 10.5194/acp-14-9363-2014, 2014.

1 Supplementary Material

1.1 CO_2 Measurements

A single calibration standard was kept at each site, and run periodically in order to assess whether any drift occurred in the CO_2 measurements over time, to determine if the calibration coefficients required any adjustment. At Robben Island the measurements of the calibration standard were (mean $\pm$ (standard deviation)) 386.89 ppm ($\pm$ 0.014) in November, 385.67 ppm ($\pm$ 0.012) in February 2013, and 385.73 ppm ($\pm$ 0.012) in June 2013. This indicates that the instrument was making stable measurements during the sixteen month campaign. The slightly higher reading in November 2012 occurred when the weather was wet, and there was more moisture contamination from the previous ambient measurements prior to calibration. At Hangklip the instrument measured the calibration standard at 378.26 ppm ($\pm$ 0.009) in November 2012 and at 378.16 ppm ($\pm$ 0.022) in June 2013, indicating no evidence of drift over the measurement period.

In addition to each site's calibration standard, a travelling reference standard was also measured at close proximity in time at all three sites, including Cape Point. This standard consisted of clean air collected at the Cape Point site. The results of two of these calibration measurements are presented in Table 1. By comparing calibration measurements with Cape Point inter-site differences are ensured to be negligible, and also links the Robben Island and Hangklip sites into the greater GAW network. The instruments at Cape Point are routinely calibrated with standards shared around the GAW network, which maintains high levels of quality control. Differences were found to be small between sites, and between calibration periods using the same standard in June 2012 and June 2013.

Table 1. Mean and standard deviation of CO_2 measurements (ppm) during calibration phase using a travelling CO_2 standard (June 2012 and June 2013) traceable to several Cape Point CO_2 laboratory standards. These in turn are linked to primary CO_2 standards maintained at the NOAA calibration laboratory in Boulder, USA.

FA01830	Site			
Date	Cape Point	Robben Island	Hangklip	Maximum inter-site difference
June 2012	452.77 (0.03)	452.88 (0.02)	452.89 (0.03)	0.12
June 2013	452.28 (0.03)	452.86 (0.02)	452.60 (0.03)	0.58
June 2012 - June 2013	0.49	0.02	0.29	

The observed hourly CO_2 concentrations at the Robben Island and Hangklip sites are presented in Figure 1, together with the hourly measurements at Cape Point and the measured daily temperature at this site. There is no clear correlation between the peaks in CO_2 and daily temperature. From March 2012 until June 2013, the mean CO_2 concentration observed at Robben Island was 391.3 ppm ($\pm$ 5.02), usually ranging between 389.5 and 394.2 ppm, with a minimum of 382.4 ppm and a maximum of 445.0 ppm. The measurements at Hangklip had a similar mean of 390.6 ppm ($\pm$ 3.89), usually ranging between 389.5 and 391.4 ppm with a minimum of 380.4 ppm and a maximum of 430.6 ppm. The Cape Point measurements have a narrower range of 382.9 to 412.3 ppm, with a mean of 392.1 ppm, indicating less influence from local sources and sinks.

The mean diurnal cycle for each month is presented in Figure 2. Across all months, the diurnal cycle of CO_2 concentrations at Cape Point are relatively flat compared with Robben Island and Hangklip. In November 2012 and February 2013 the diurnal cycle for both measurement sites was the most flat. This is the summer period, when temperatures are high and the Western Cape experiences the lowest amount of rain. The amplitude of the diurnal cycle at both sites increased from April, reaching a maximum amplitude in June and July. This is during the winter rainfall period in the Western Cape. Temperatures are mild and much of the vegetation growth occurs during this period. The diurnal cycle of the Hangklip site dipped below both Cape Point and Robben Island, indicating that this site is more affected by local sinks of CO_2. Robben Island consistently had the highest peaks in CO_2 concentrations across all months, indicating that this site was the most affected by local sources, which is what we expected.

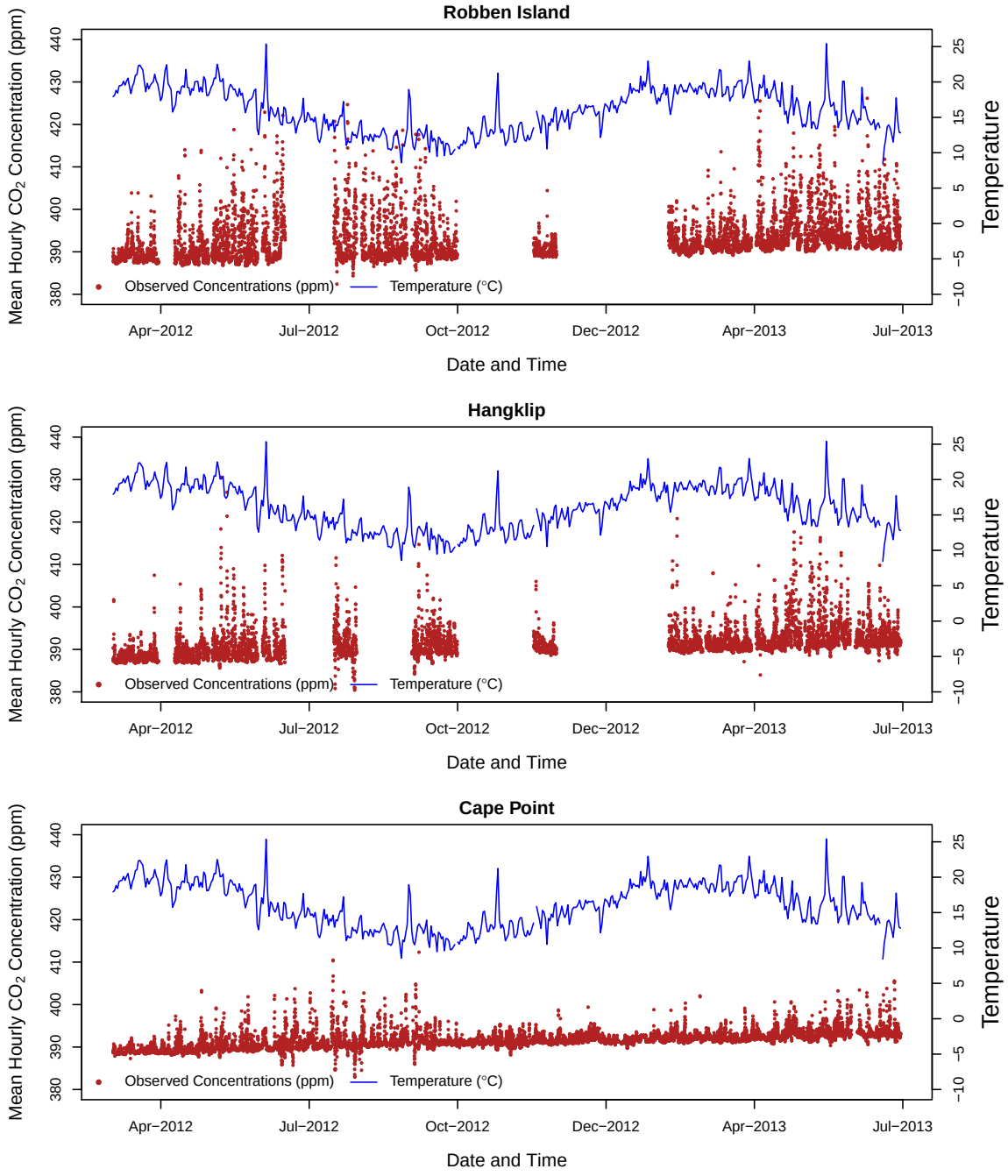


Figure 1. Observed hourly CO_2 concentrations (ppm) (left-side axis) at the Robben Island (top closed red circles) and Hangklip (bottom closed black circles) measurement sites. The blue line appearing at the bottom of each plot is the CO_2 concentration measurements at Cape Point station (ppm) and the green line at the top of each plot is the mean daily temperature $(^{\circ}\mathrm{C})$ as measured at the Cape Point station, which is represented by the right-side axis.

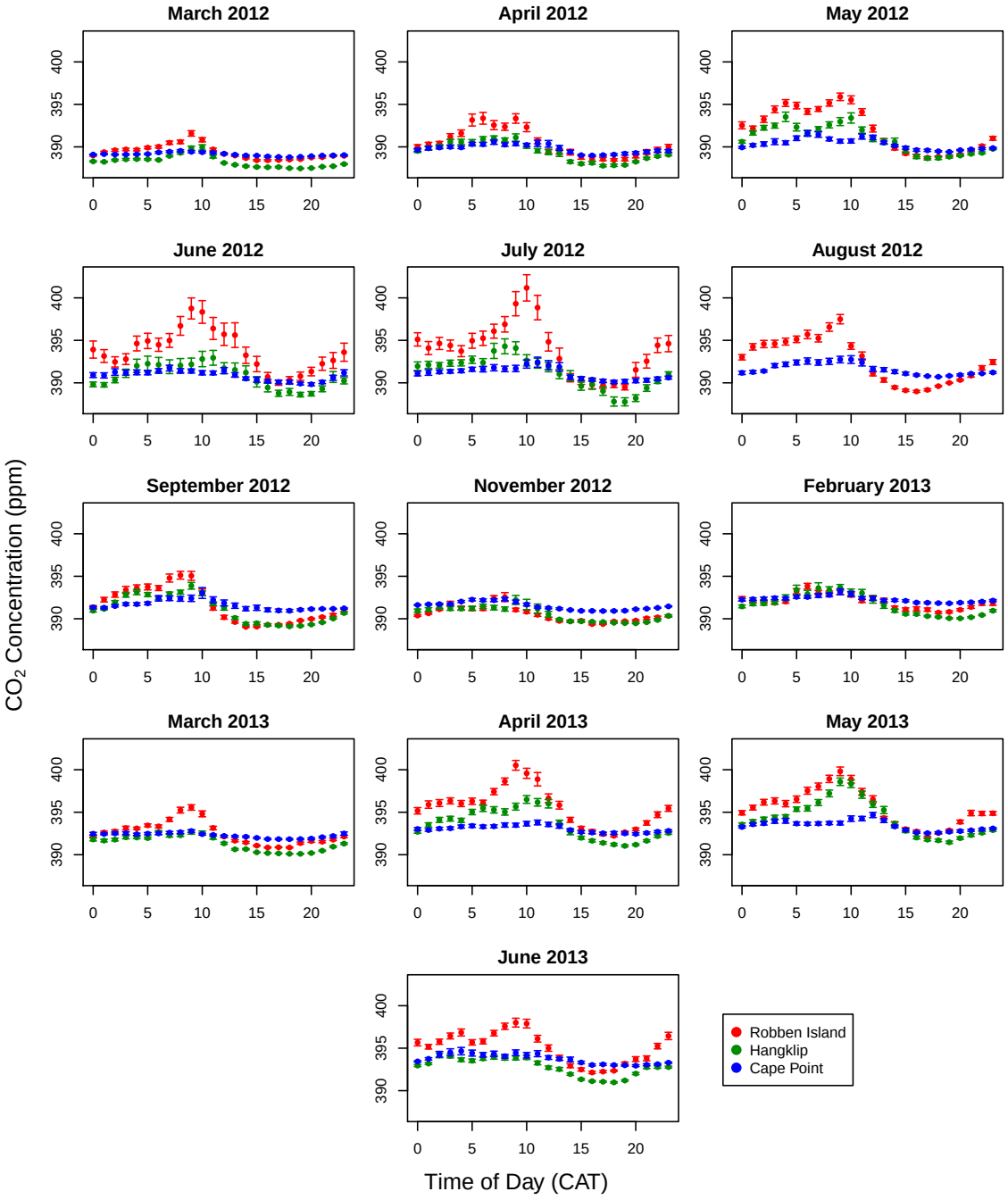


Figure 2. Diurnal cycle of the observed CO_2 concentrations (ppm) for each month and at each site with 95% confidence intervals, where the standard error is calculated over all measurements available for that hour of the day during that particular month. Cape Point is the generally flat diurnal cycle in blue, Robben Island with the generally larger daytime CO_2 concentrations in red, and Hangklip with the generally lower afternoon CO_2 concentrations in black.

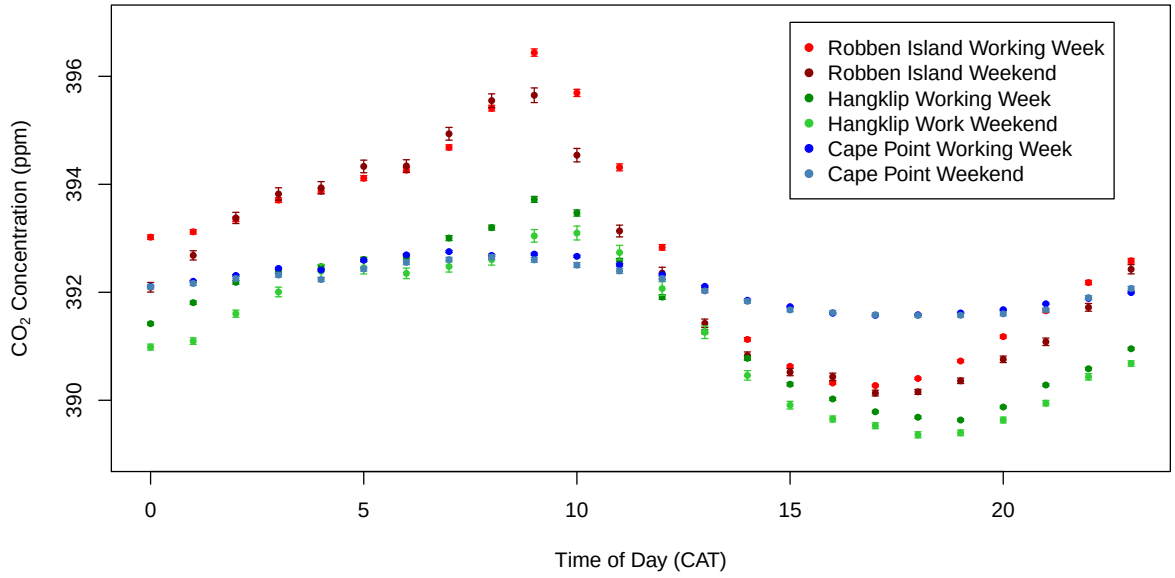


Figure 3. Diurnal cycle of the observed CO_2 concentrations (ppm) over the full measurement period from March 2012 until June 2013 at each site with 95% confidence intervals, where the standard error is calculated over all measurements available for that hour of the day during the entire measurement period, separated by site (Cape Point - blue and light blue, Robben Island - red and dark red, Hangklip - black and grey) and by working week (brighter colour) and weekend (duller colour).

The mean diurnal cycle over the whole measurement period is presented in Figure 3, separated by site and by working week and weekend. The background site, Cape Point, shows no discernible difference between the mean concentrations over the week and weekend, whereas Robben Island and Hangklip sites measure concentrations during the working week which tend to be larger across most of the day. Both the early morning and afternoon means show a clear tendency for these sites to have larger concentrations during the working week compared with the same time of day over the weekend, which can only be due to anthropogenic influences. This supports the separation of fossil fuel fluxes into working week and weekend contributions.

Table 2. The number of days with available CO_2 measurement data for each month (out of a possible four week period considered) and overall percentage available data out of 16 four-week periods for each site.

Site	Year / Month																Overall Percentage
	2012										2013						
	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	
Robben Island	28	21	28	14	14	28	28	0	14	0	0	21	28	28	28	28	68.75
Hangklip	28	21	28	14	14	0	28	0	14	0	0	21	28	28	28	28	62.50
Combined	56	42	56	28	28	28	56	0	28	0	0	42	56	56	56	56	65.63

Table 2 provides the measured CO_2 concentration availability at each site. Robben Island had slightly higher data availability compared with Hangklip, with a 65.6% data availability overall for the sixteen month measurement campaign. Each site was equipped with a Hauwei USB modem connected to a 3G network, and were set to submit data to an email address on a regular basis. Through these emails or through connecting remotely to the instruments, instrumentation problems could be detected. Most of the down-time at the sites was attributed to either pump failure, or occasionally the instrument software had failed and there was no available person to restart the instrument due to limited access to these sites. Robben Island lighthouse is manned, and therefore it was possible to request the lighthouse keeper to restart the instrument when frozen, but more regular than expected pump problems required visits to the site to replace the offending device. This entailed making arrangements for unplanned voyages to Robben Island, which could take some time to arrange. Hangklip is unmanned, and the site has strict access control, therefore problems at this site tended to take slightly longer to remedy. The final four months considered in this study had the best continuity of data availability.

1.2 Model Assessment

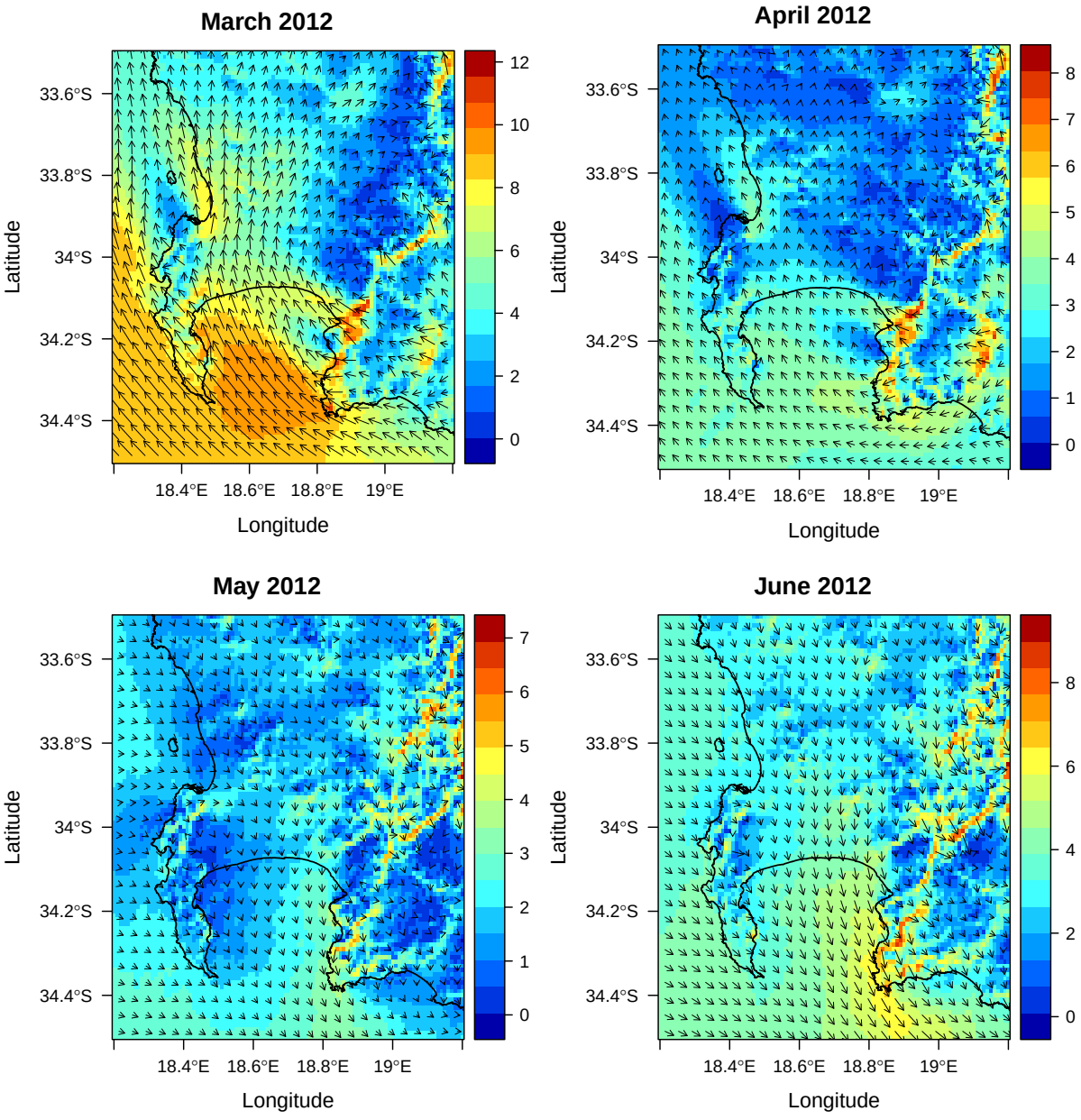
To determine if the prescribed prior covariance parameters were consistent with the model assumptions, the sum of the squared normalised residuals was compared against the χ^2 distribution. For most months the standardised statistic was close to one, but in the case of June and July 2012, this statistic was above 2. We did not scale the variances independently for each month, and therefore the single scaling factor of 2 for the prior flux variances was not large enough for all months. The statistic remained below 2.5 for all months, and had a minimum of 0.68 for the month of November 2012. The mean of the statistic over all months was 1.48. A subsequent study will assess an alternative approach to determine prior flux uncertainties, which would guarantee compliance of the sum of the squared normalised residuals to follow the χ^2 distribution.

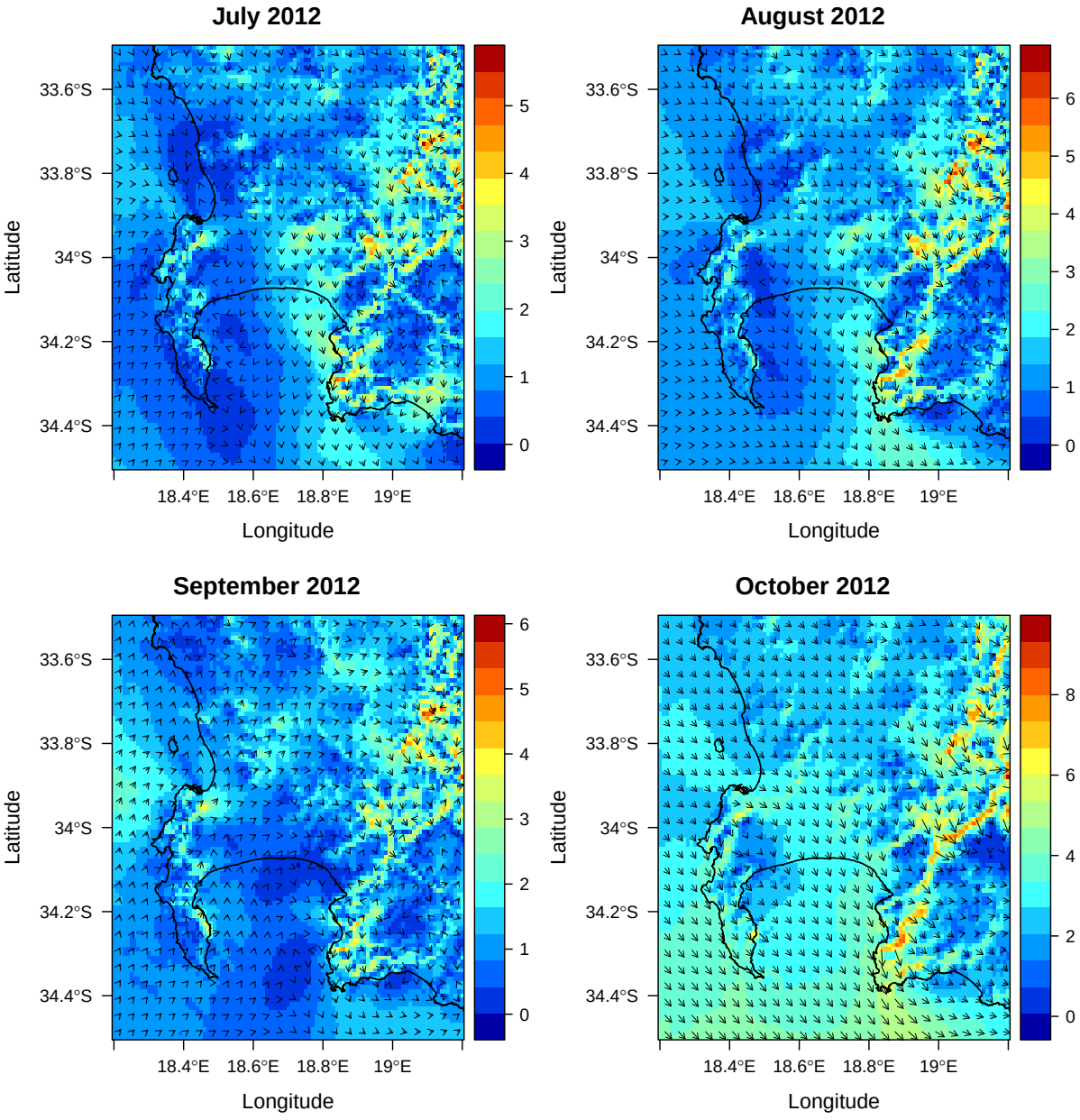
Table 3. Sum of the squared normalised residuals for each month, which should approximate a χ^2 distribution with one degree of freedom.

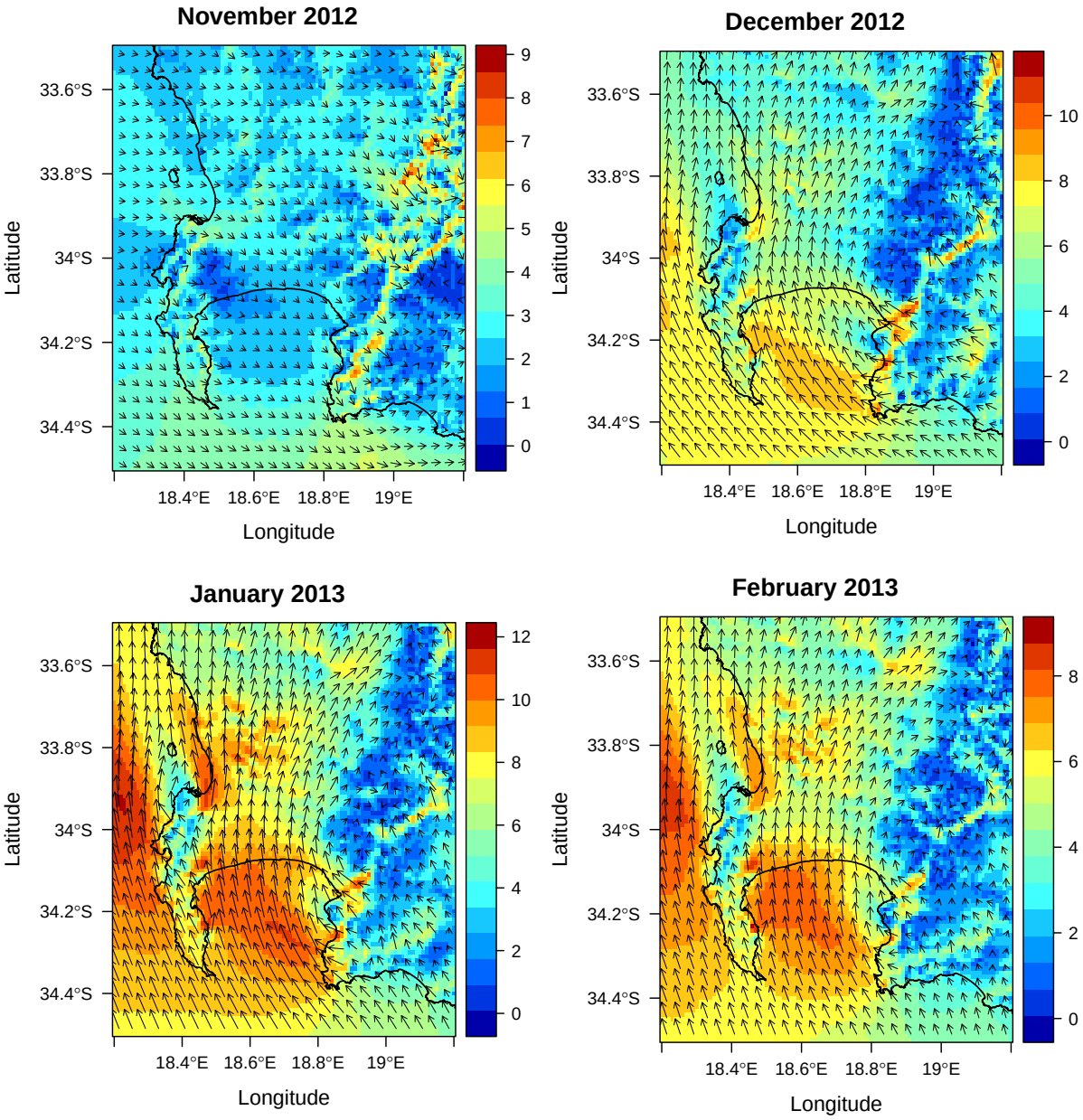
Month	χ^2 statistic	Percentage available data
Mar2012	1.006	100%
Apr2012	1.089	75%
May2012	1.959	100%
Jun2012	2.473	50%
Jul2012	2.170	50%
Aug2012	1.586	50%
Sep2012	1.028	100%
Nov2012	0.678	50%
Feb2012	0.938	75%
Mar2013	1.048	100%
Apr2013	1.776	100%
May2013	1.829	100%
Jun2013	1.706	100%

1.3 Modelled Wind

Figure 4 provides the average wind speed and direction for the domain for each month, as modelled by CCAM. The plots were created using the R package rasterVis (Lamigueiro, 2016).




July 2012
August 2012
September 2012
October 2012
Latitude
Longitude
33.6°S
33.8°S
34°S
34.2°S
34.4°S
18.4°E
18.6°E
18.8°E
19°E


November 2012
December 2012
January 2013
February 2013
Latitude
Longitude
33.6°S
33.8°S
34°S
34.2°S
34.4°S
18.4°E
18.6°E
18.8°E
19°E

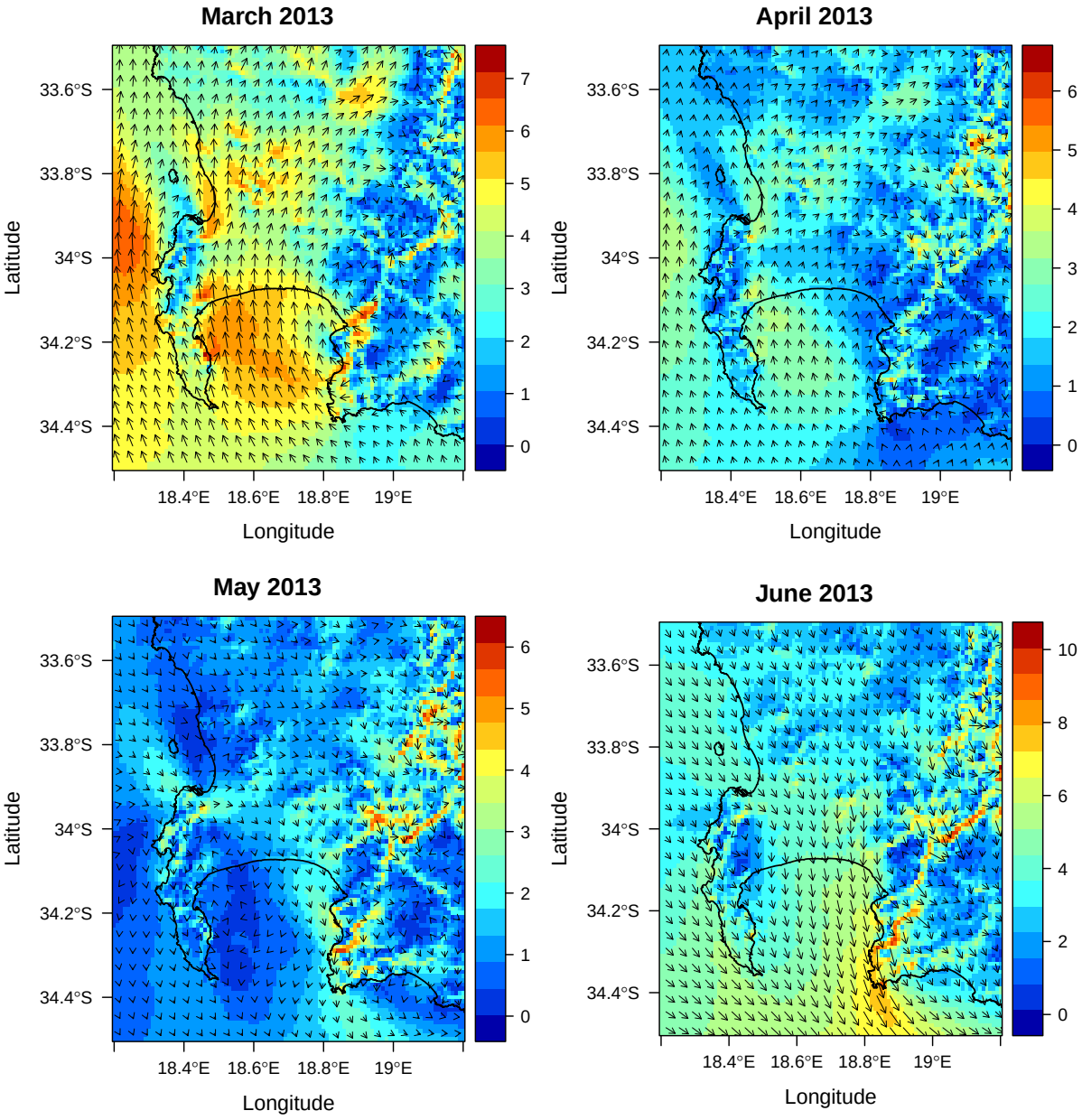
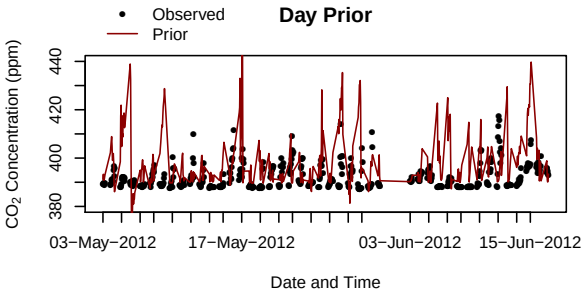
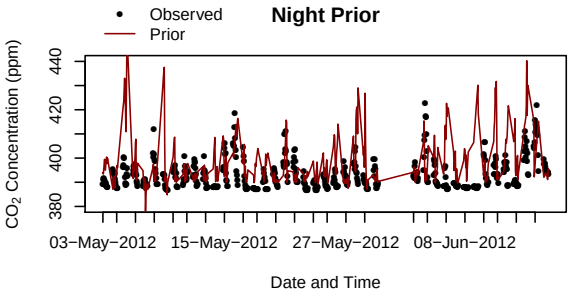


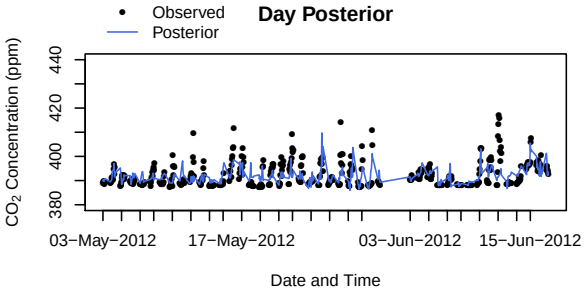
Figure 4. Mean modelled wind speed and direction in the Cape Town domain for each month. The colourbar represents the mean wind speed (m/s).

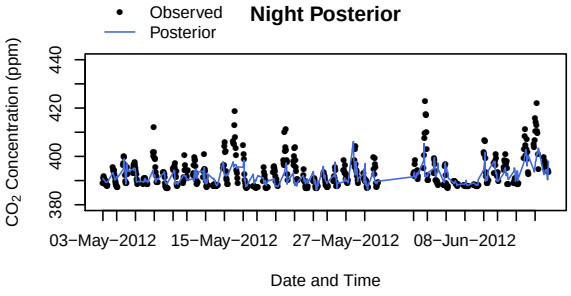
1.4 Modelled Concentrations

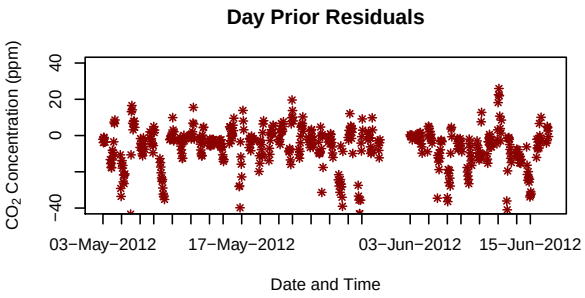
The time series for the modelled concentrations is discussed in Section 3.1 of the main paper.

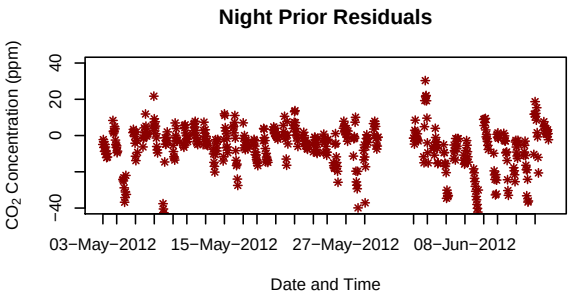

Observed
Prior
Day Prior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
17−May−2012
03−Jun−2012
15−Jun−2012
Date and Time

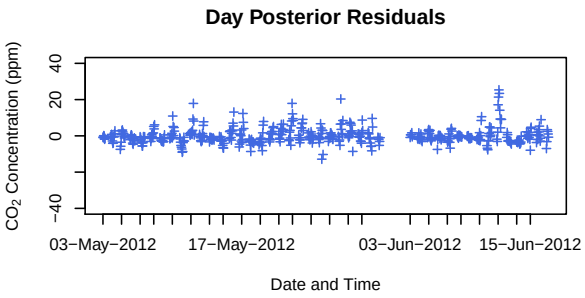

Observed
Prior
Night Prior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
15−May−2012
27−May−2012
08−Jun−2012
Date and Time

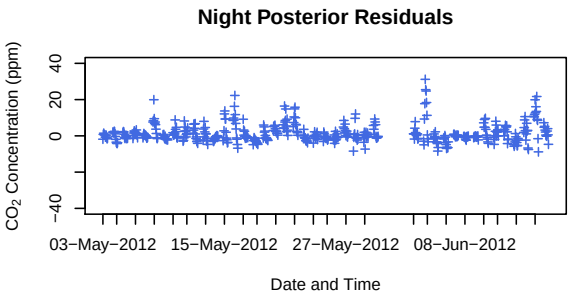

Observed
Posterior
Day Posterior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
17−May−2012
03−Jun−2012
15−Jun−2012
Date and Time

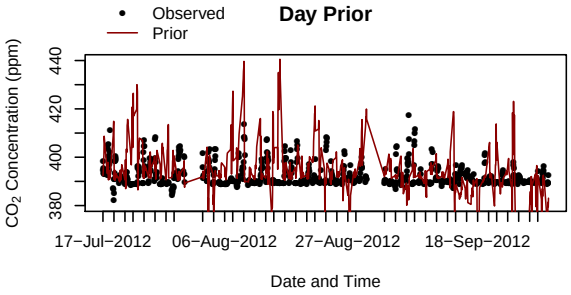

Observed
Posterior
Night Posterior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
15−May−2012
27−May−2012
08−Jun−2012
Date and Time

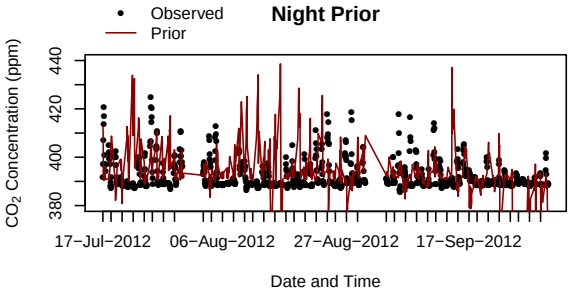

Day Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
17−May−2012
03−Jun−2012
15−Jun−2012
Date and Time

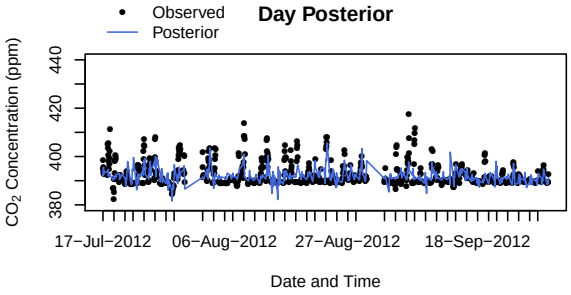

Night Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
15−May−2012
27−May−2012
08−Jun−2012
Date and Time

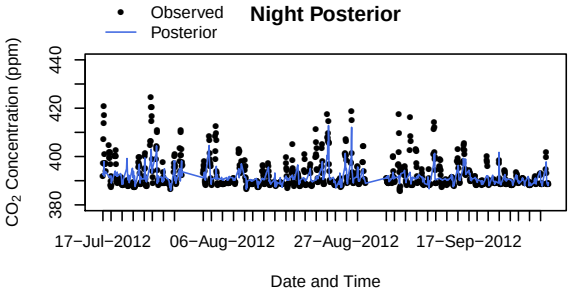

Day Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
17−May−2012
03−Jun−2012
15−Jun−2012
Date and Time

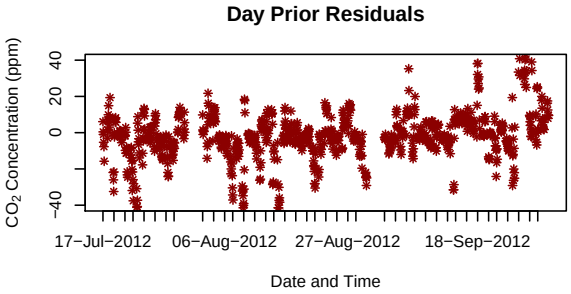

Night Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
15−May−2012
27−May−2012
08−Jun−2012
Date and Time

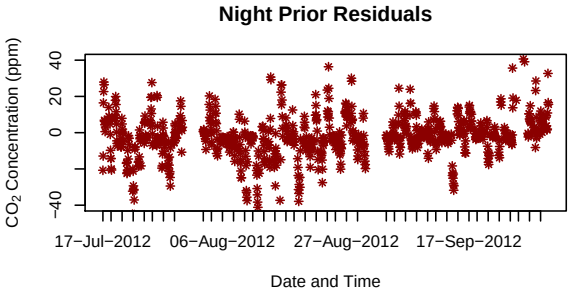

Observed
Prior
Day Prior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
06−Aug−2012
27−Aug−2012
18−Sep−2012
Date and Time

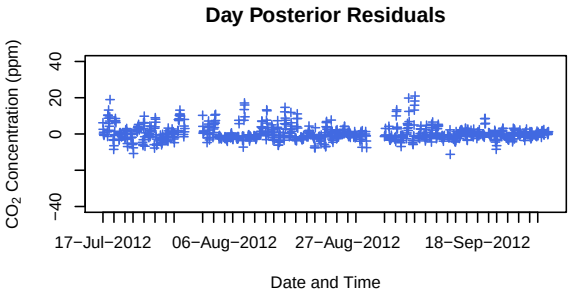

Observed
Prior
Night Prior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
06−Aug−2012
27−Aug−2012
17−Sep−2012
Date and Time

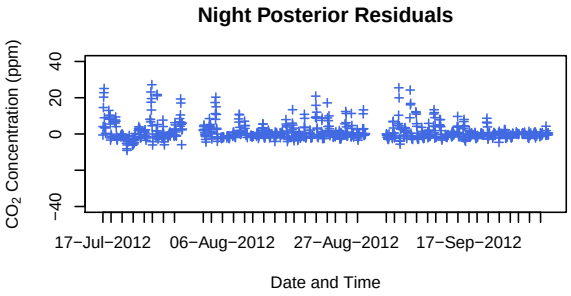

Observed
Posterior
Day Posterior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
06−Aug−2012
27−Aug−2012
18−Sep−2012
Date and Time

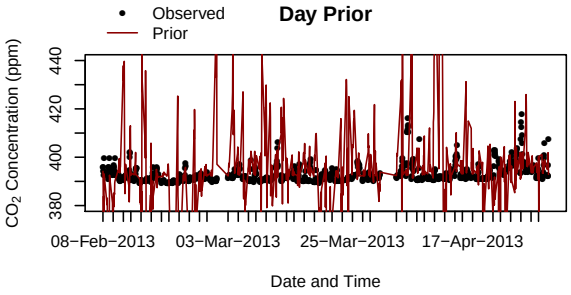

Observed
Posterior
Night Posterior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
06−Aug−2012
27−Aug−2012
17−Sep−2012
Date and Time

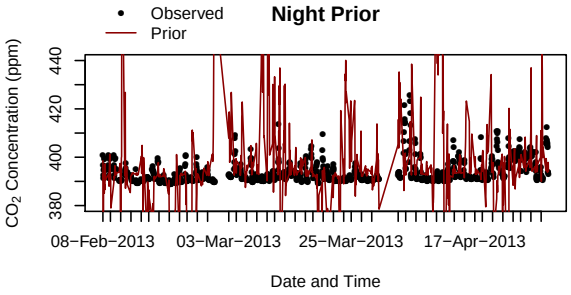

Day Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
06−Aug−2012
27−Aug−2012
18−Sep−2012
Date and Time

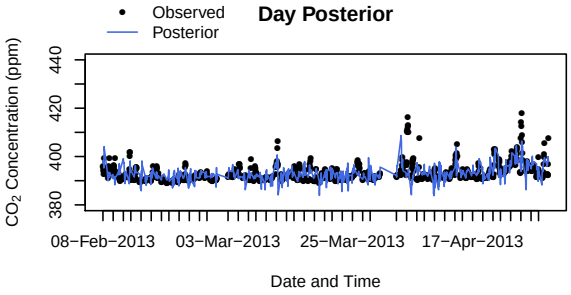

Night Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
06−Aug−2012
27−Aug−2012
17−Sep−2012
Date and Time

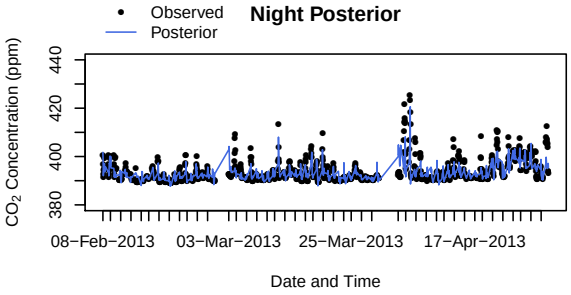

Day Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
06−Aug−2012
27−Aug−2012
18−Sep−2012
Date and Time

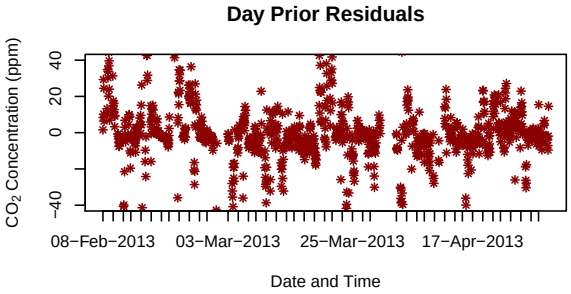

Night Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
06−Aug−2012
27−Aug−2012
17−Sep−2012
Date and Time

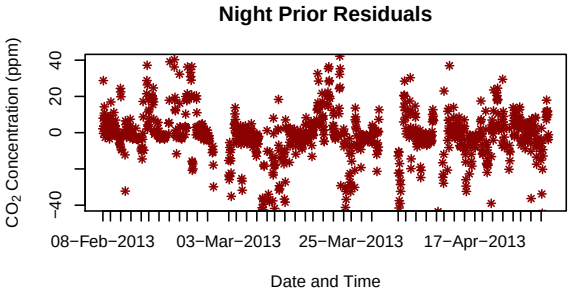

Observed
Prior
Day Prior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time

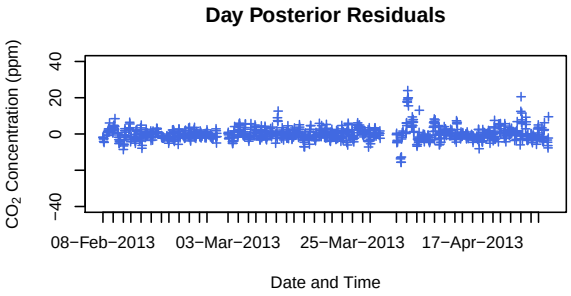

Observed
Prior
Night Prior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time

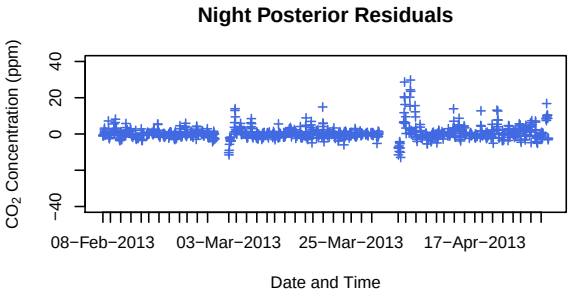

Observed
Posterior
Day Posterior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time


Observed
Posterior
Night Posterior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time


Day Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time


Night Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time


Day Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time


Night Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
03−Mar−2013
25−Mar−2013
17−Apr−2013
Date and Time

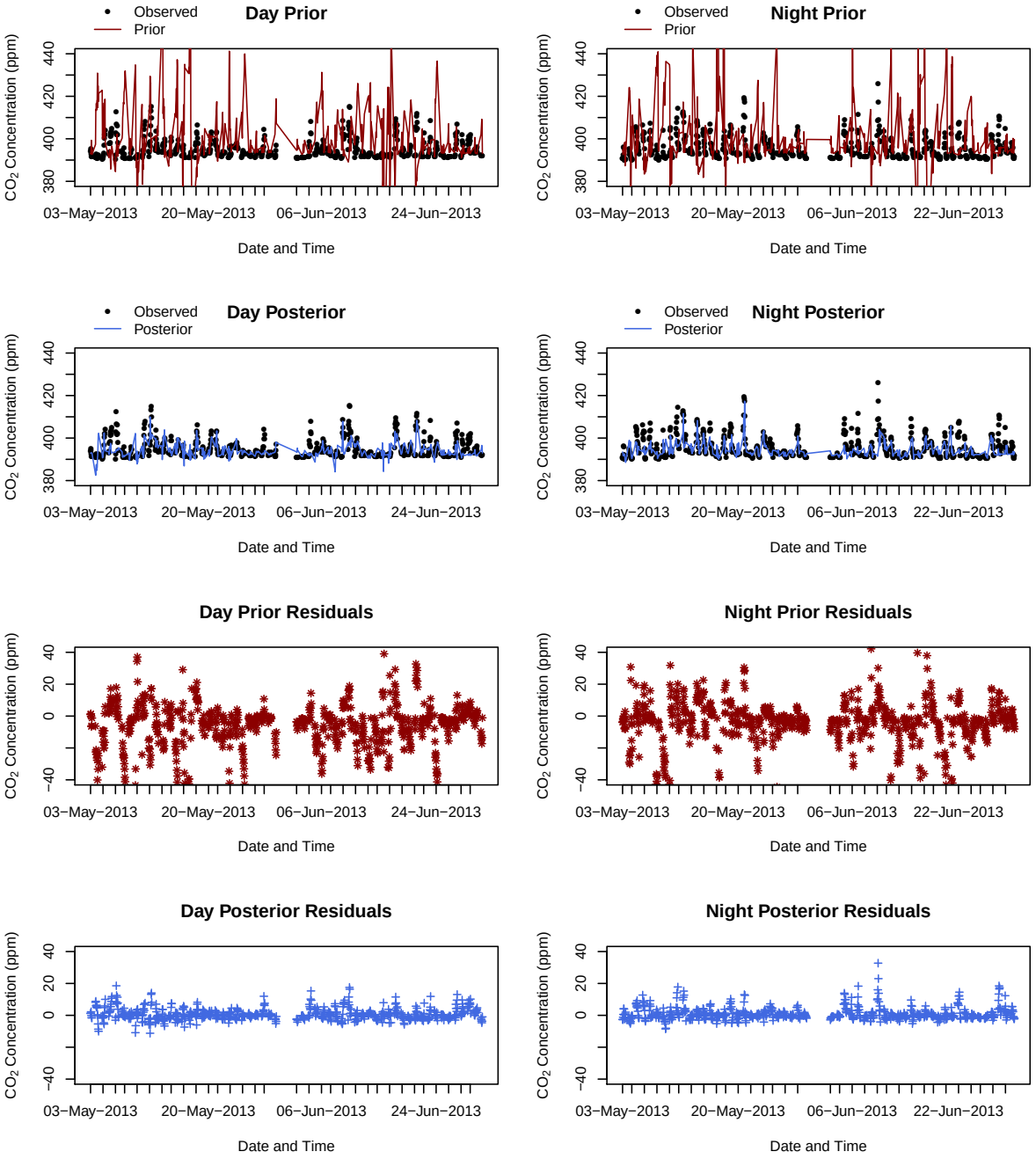
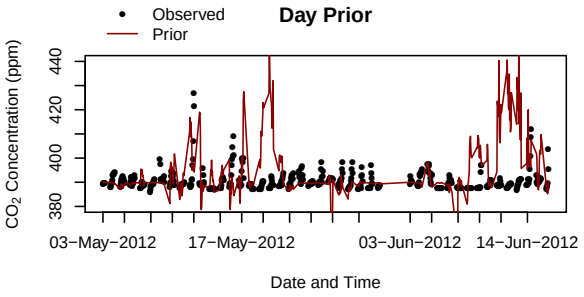
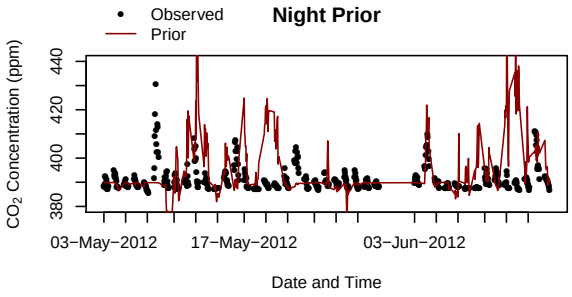
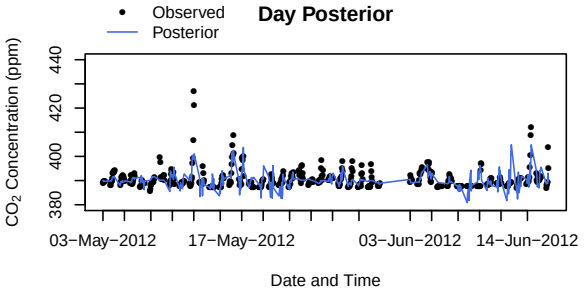
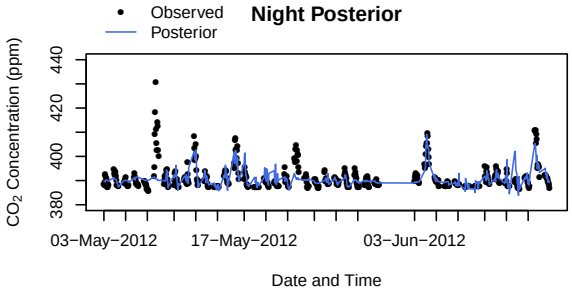


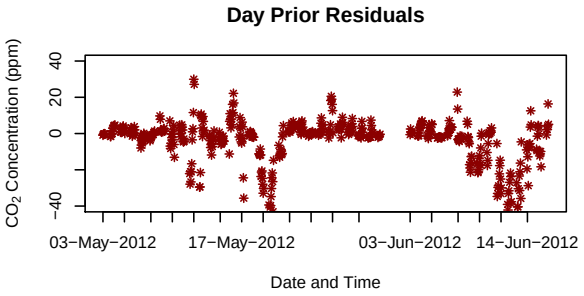
Figure 5. The top 4 panels provide a time series of the observed, prior and posterior modelled concentrations at the Robben Island site. The time series is separated into day and night-time periods. The residuals between the modelled and observed concentrations, defined as the difference between the observed and modelled concentrations, are provided in the lower panel 4 panels. The first two months are presented in the main paper in section 3.1.

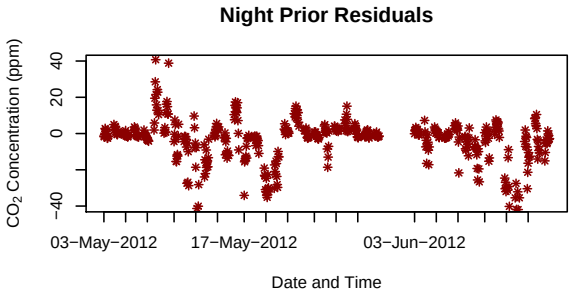

Observed
Prior
Day Prior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
17−May−2012
03−Jun−2012
14−Jun−2012
Date and Time

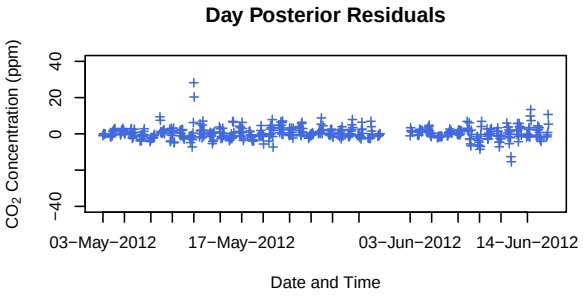

Observed
Prior
Night Prior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
17−May−2012
03−Jun−2012
Date and Time

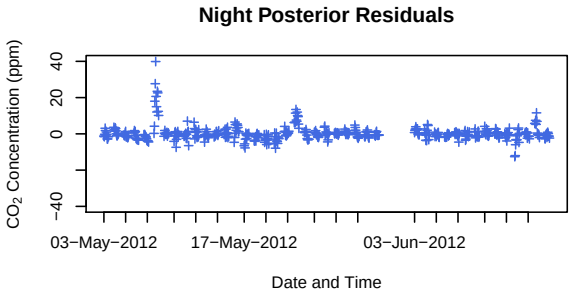

Observed
Posterior
Day Posterior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
17−May−2012
03−Jun−2012
14−Jun−2012
Date and Time

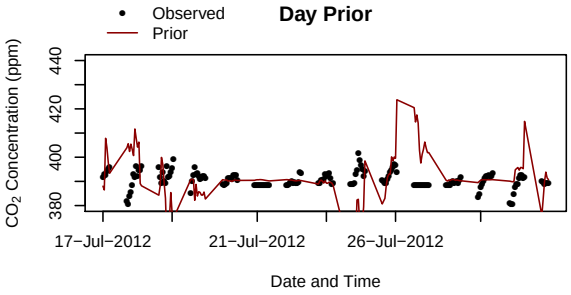

Observed
Posterior
Night Posterior
CO2 Concentration (ppm)
380
400
420
440
03−May−2012
17−May−2012
03−Jun−2012
Date and Time

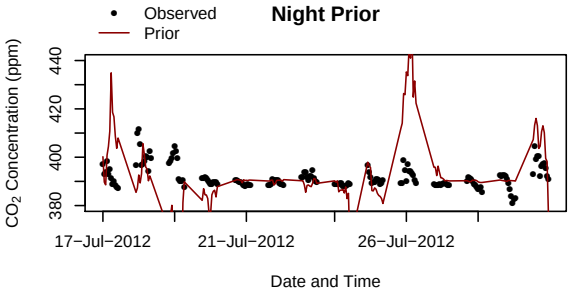

Day Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
17−May−2012
03−Jun−2012
14−Jun−2012
Date and Time

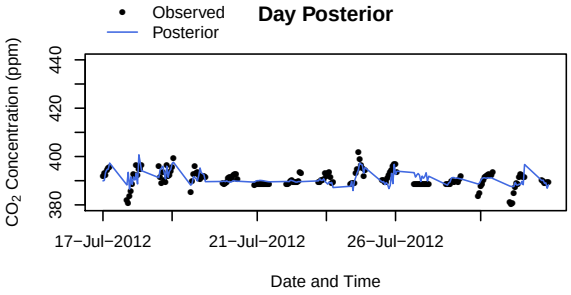

Night Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
17−May−2012
03−Jun−2012
Date and Time

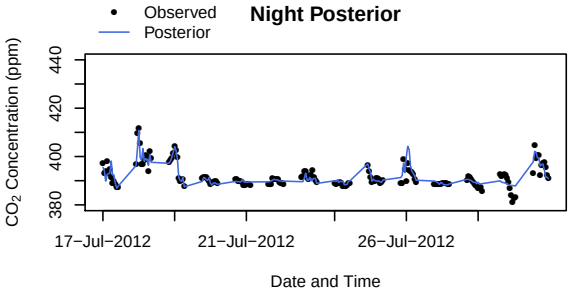

Day Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
17−May−2012
03−Jun−2012
14−Jun−2012
Date and Time

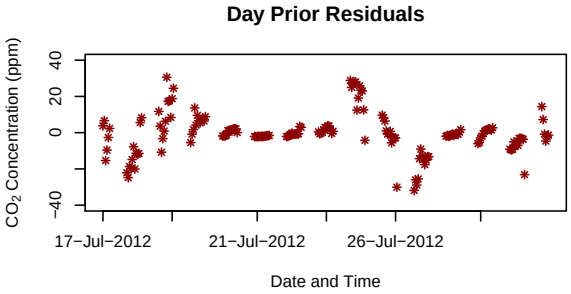

Night Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
03−May−2012
17−May−2012
03−Jun−2012
Date and Time

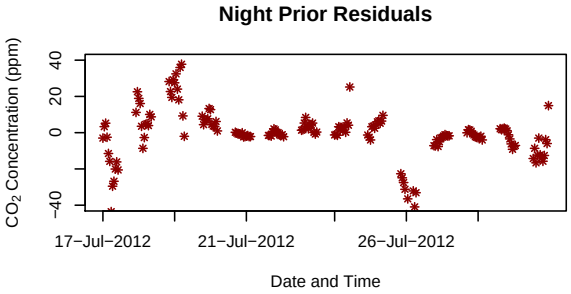

Observed
Prior
Day Prior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

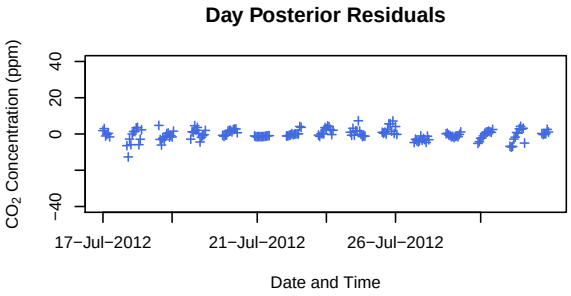

Observed
Prior
Night Prior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

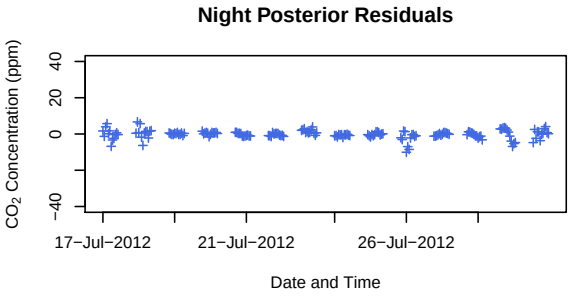

Observed
Posterior
Day Posterior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

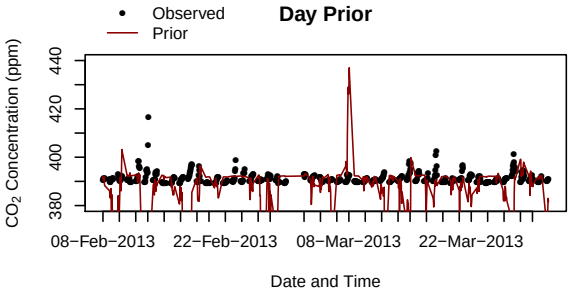

Observed
Posterior
Night Posterior
CO2 Concentration (ppm)
380
400
420
440
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

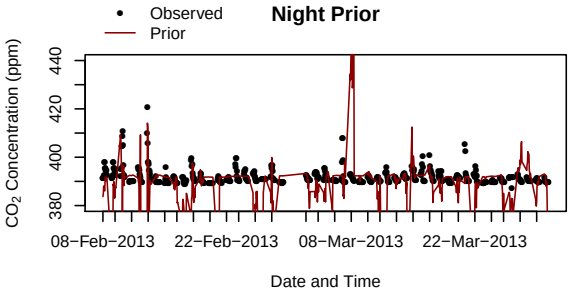

Day Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

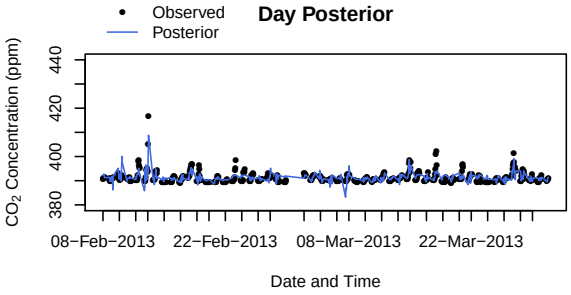

Night Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

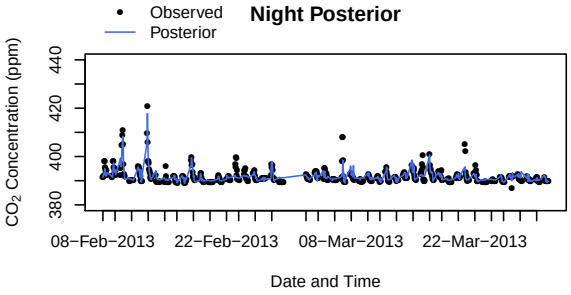

Day Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

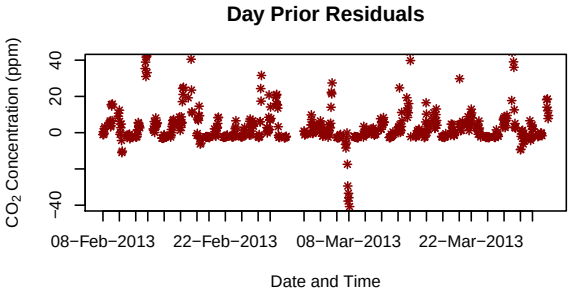

Night Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
17−Jul−2012
21−Jul−2012
26−Jul−2012
Date and Time

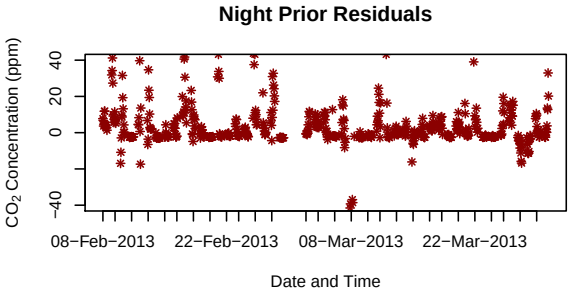

Observed
Prior
Day Prior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time

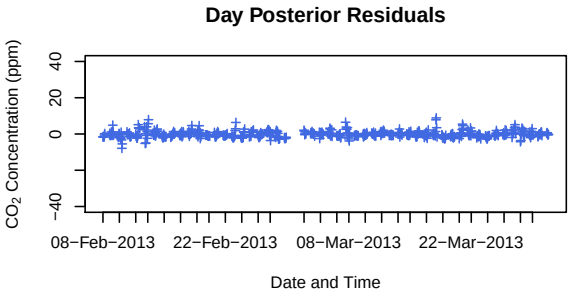

Observed
Prior
Night Prior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time

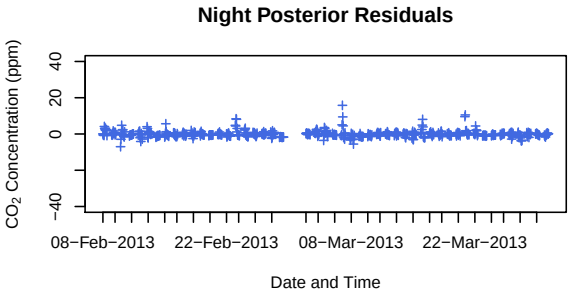

Observed
Posterior
Day Posterior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time


Observed
Posterior
Night Posterior
CO2 Concentration (ppm)
380
400
420
440
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time


Day Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time


Night Prior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time


Day Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time


Night Posterior Residuals
CO2 Concentration (ppm)
−40
0
20
40
08−Feb−2013
22−Feb−2013
08−Mar−2013
22−Mar−2013
Date and Time

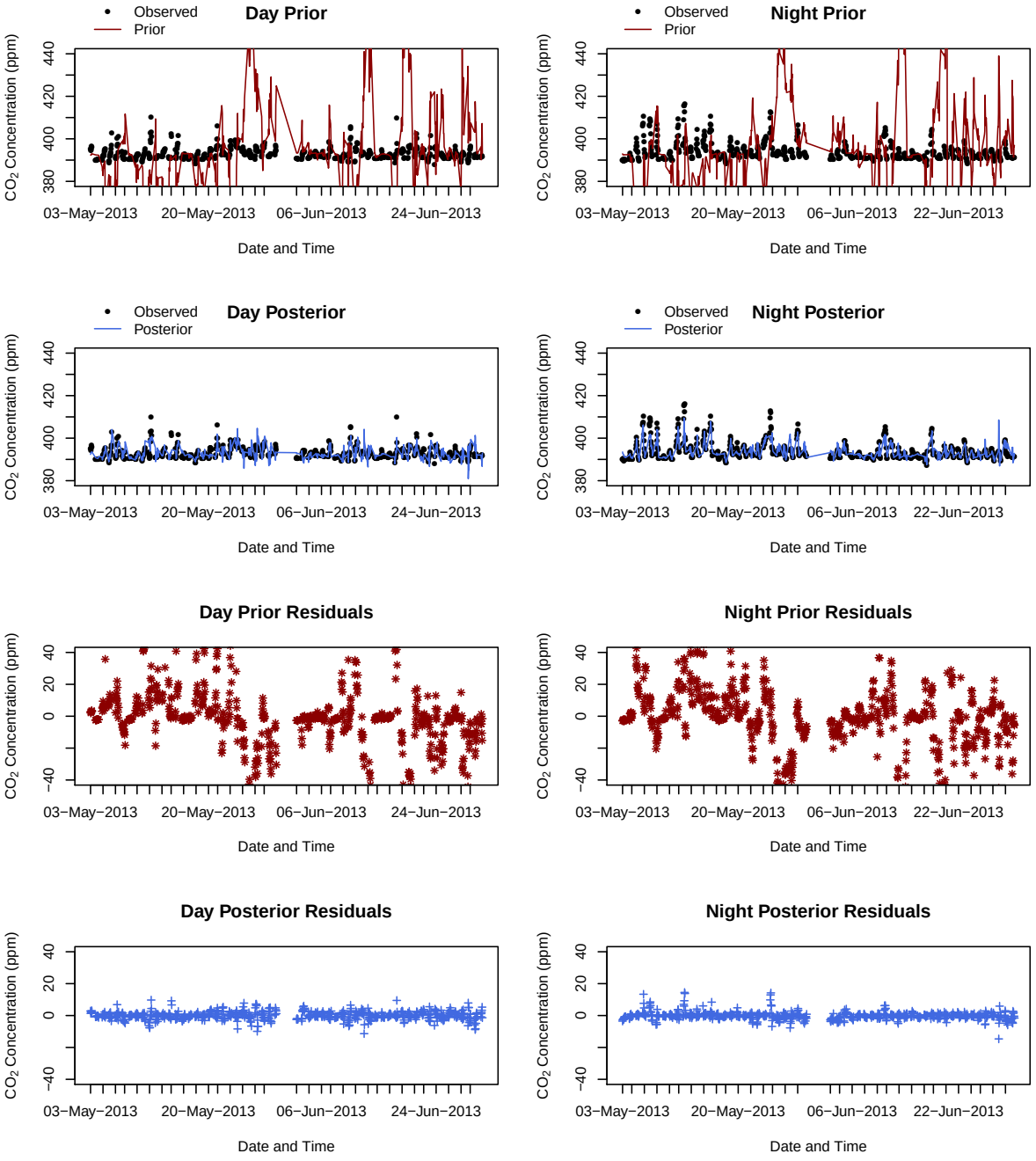


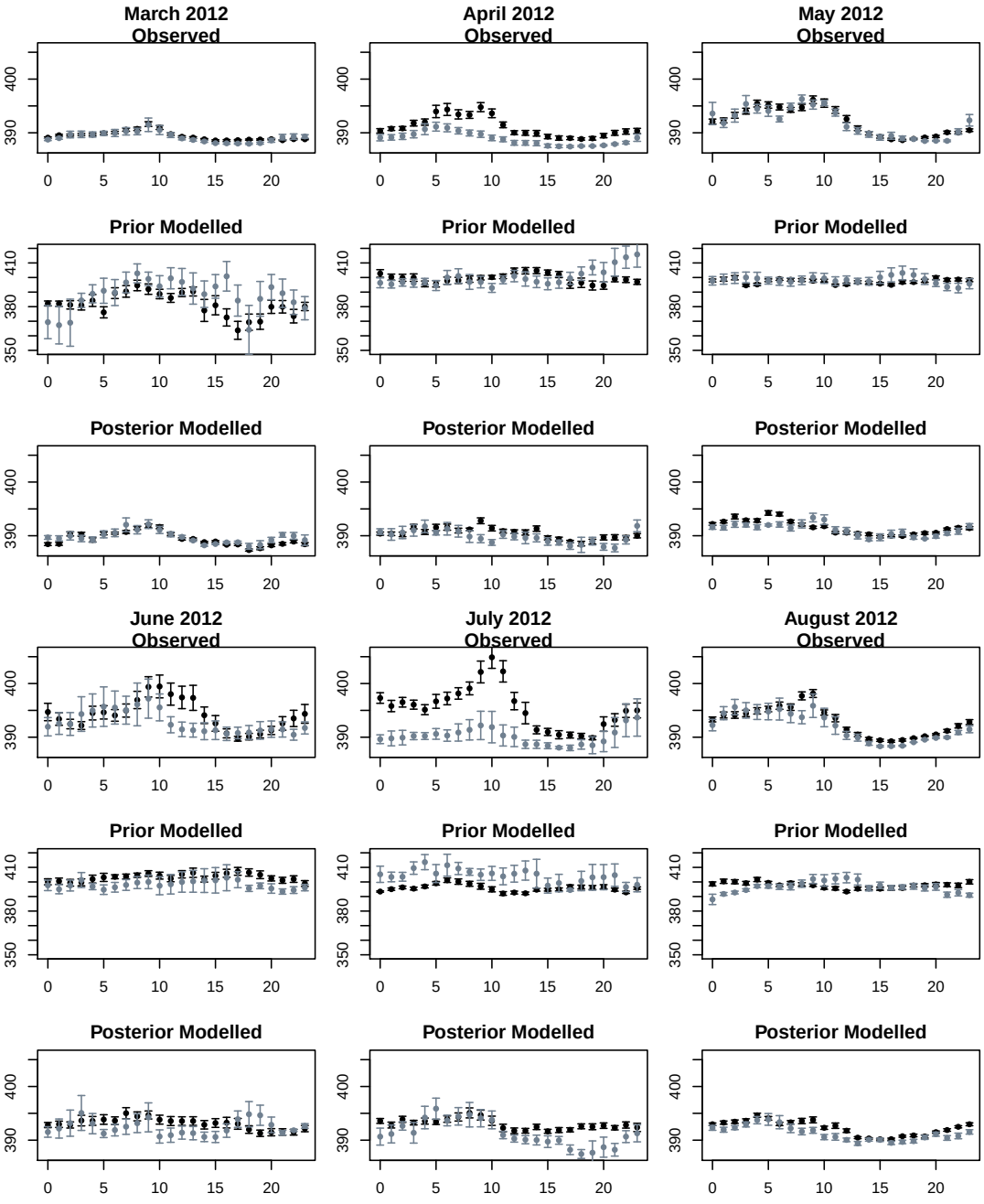
Figure 6. The top 4 panels provide a time series of the observed, prior and posterior modelled concentrations at the Hangklip site. The time series is separated into day and night-time periods. The residuals between the modelled and observed concentrations, defined as the difference between the observed and modelled concentrations, are provided in the lower panel 4 panels. The first two months are presented in the main paper in section 3.1.

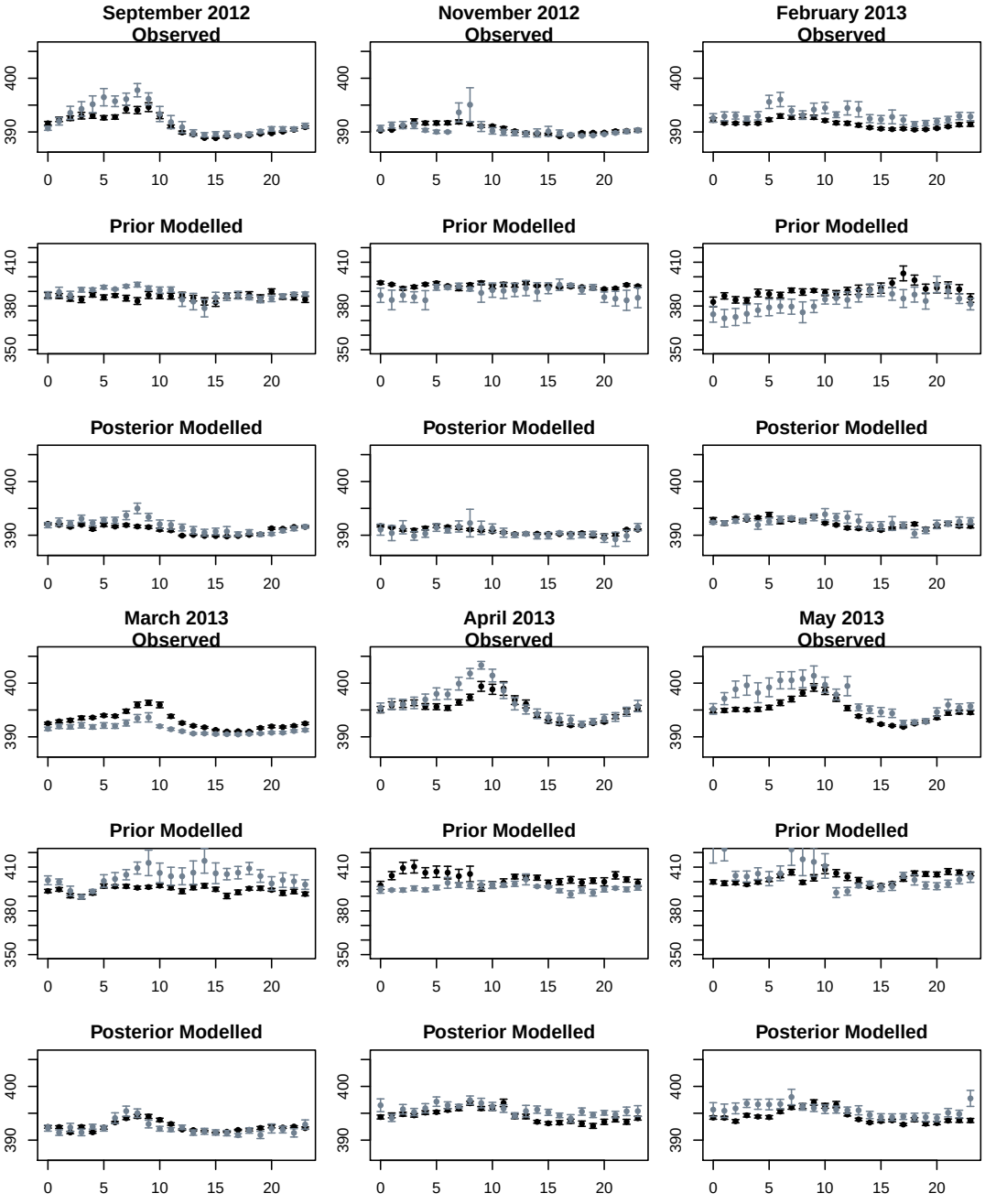
1.5 Diurnal Cycle

The observed, prior and posterior modelled diurnal cycle, separated into working week and weekend CO_2 concentrations, are provided for each site and for each month in Figures 7 and 8. For all months, the diurnal cycle of the posterior modelled concentrations is relatively flat in comparison with the observed diurnal cycle, and usually sits at a higher mean level in the case of Robben Island, and at a lower mean level in the case of Hangklip. Compared with the prior modelled concentrations, the posterior diurnal cycle matches better with the observed concentrations in terms of the peaks and troughs of the cycle and in terms of the mean level of the concentrations at each hour, although the posterior cycle still appears relatively flat in comparison to the observed cycle.

April 2013 at the Robben Island site provides an example where the prior modelled concentrations had working week concentrations that were above those for the weekend during the early morning hours, whereas the observed concentrations showed the opposite situation. After the inversion, the posterior estimates had mean concentrations for the weekend that were above those for the working week during the early morning hours, matching better with the observed diurnal cycle.

Therefore the inversion does show an ability to improve estimates of the diurnal cycle, despite only separating the sources into day and night sources over a week period, and further separating the fossil fuel sources into weekend and week sources.


March 2012
Observed
April 2012
Observed
May 2012
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled
June 2012
Observed
July 2012
Observed
August 2012
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled


September 2012
Observed
November 2012
Observed
February 2013
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled
March 2013
Observed
April 2013
Observed
May 2013
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled

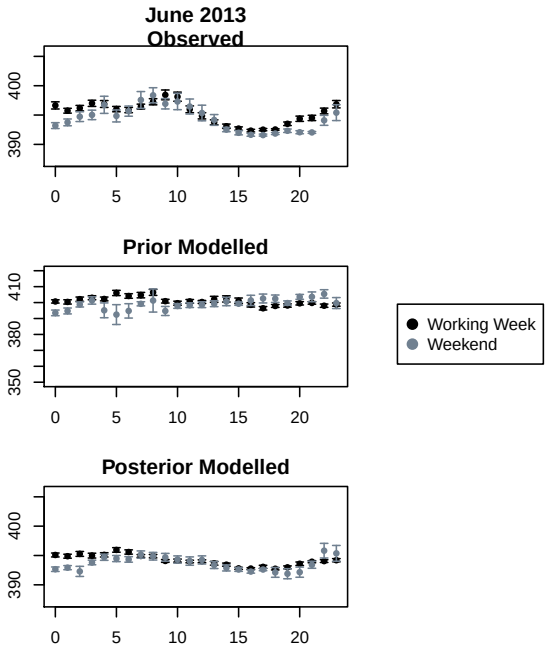
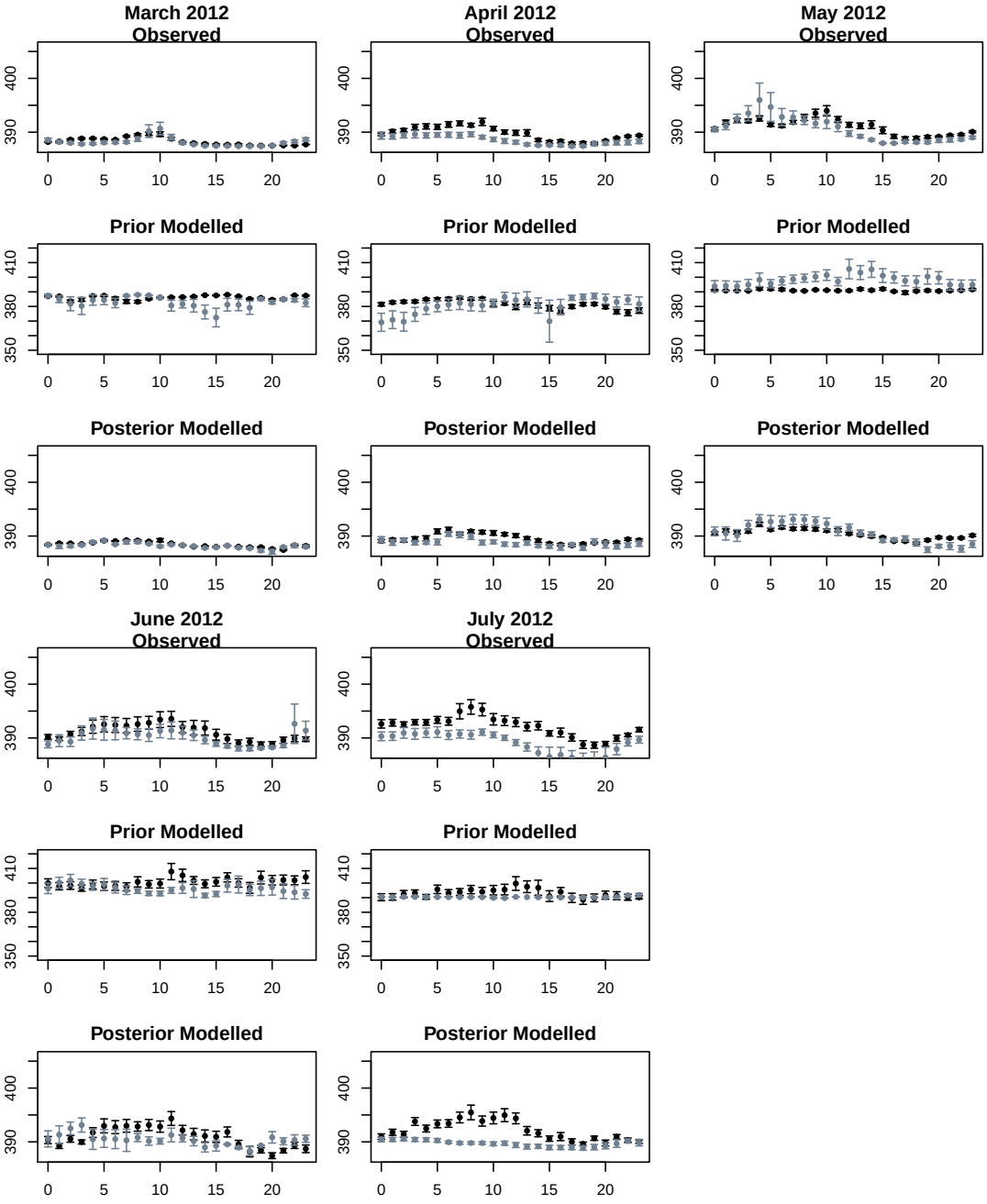
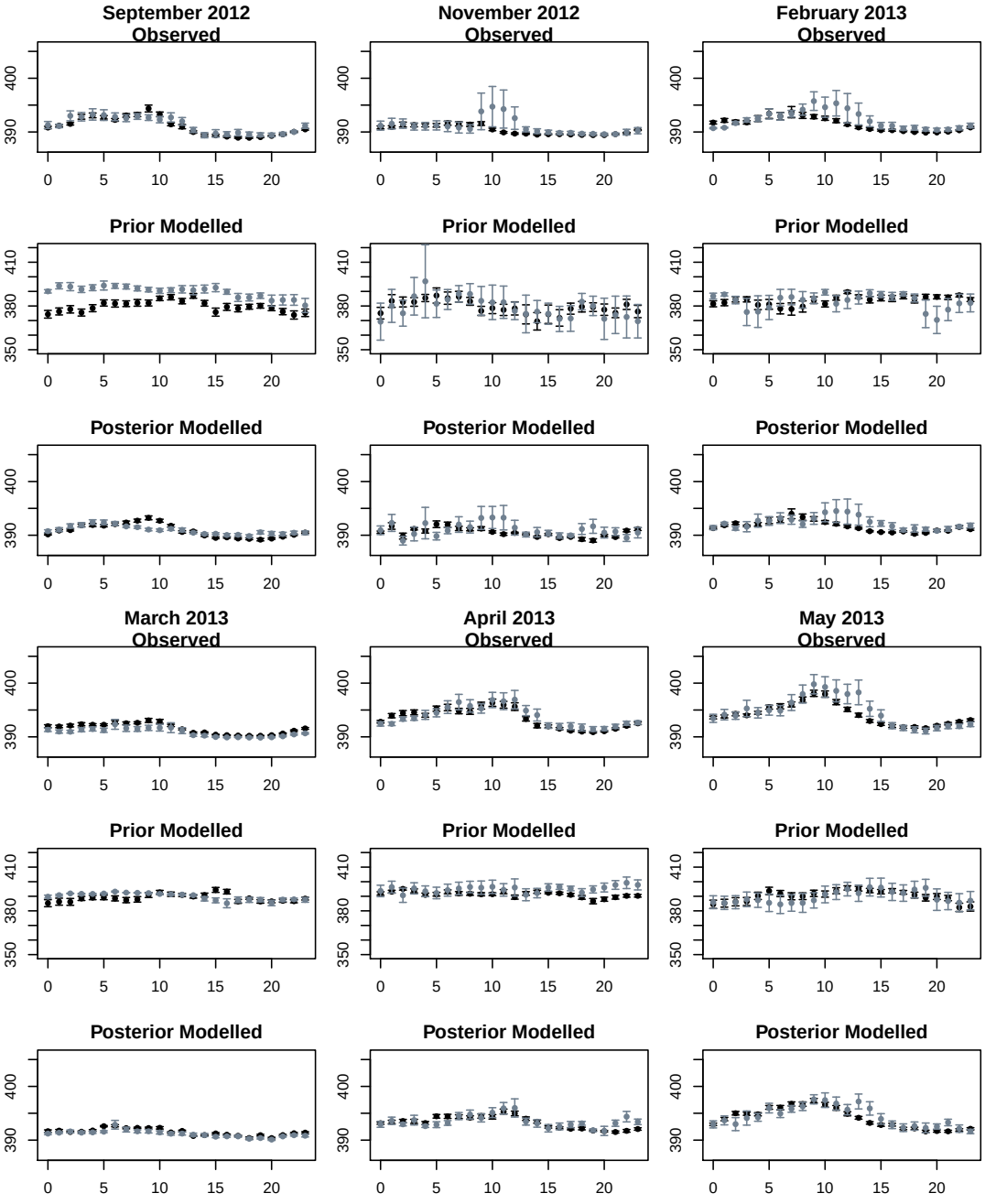


Figure 7. Diurnal cycle of the observed, prior modelled and posterior modelled CO_2 concentrations (ppm) at Robben Island, separated into working week (black) and weekend concentrations (grey), for each month with 95% confidence intervals, where the standard error is calculated over all measurements available for that hour of the day during that particular month.


March 2012
Observed
April 2012
Observed
May 2012
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled
June 2012
Observed
July 2012
Observed
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled


September 2012
Observed
November 2012
Observed
February 2013
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled
March 2013
Observed
April 2013
Observed
May 2013
Observed
Prior Modelled
Prior Modelled
Prior Modelled
Posterior Modelled
Posterior Modelled
Posterior Modelled

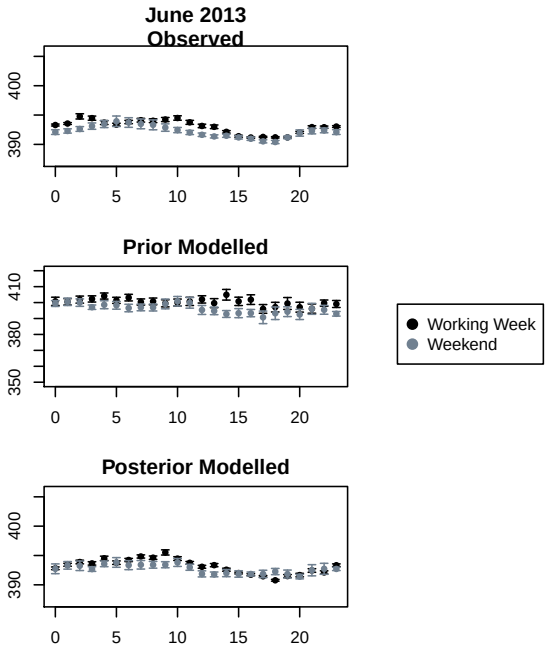
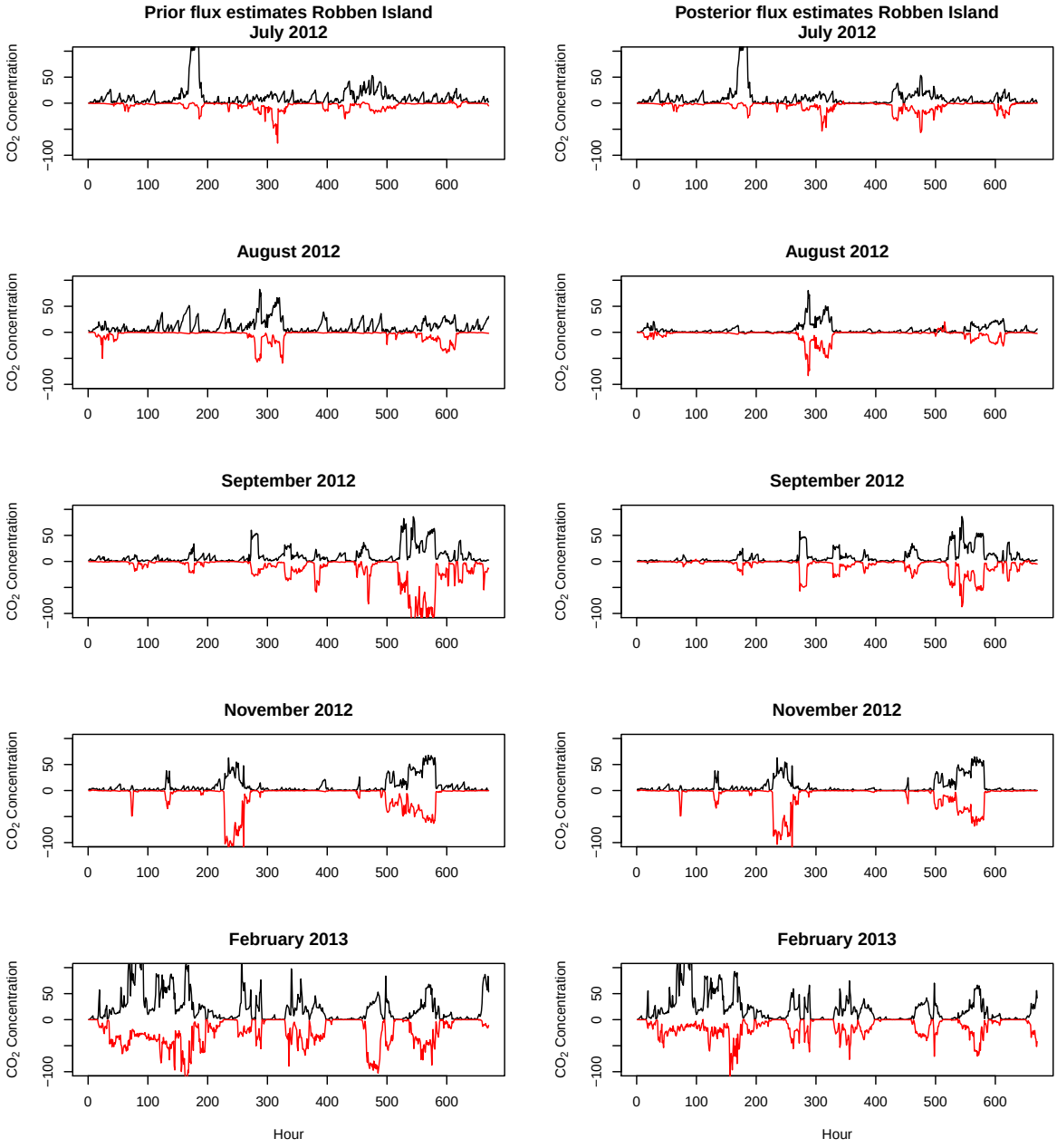


Figure 8. Diurnal cycle of the observed, prior modelled and posterior modelled CO_2 concentrations (ppm) at Hangklip, separated into working week (black) and weekend concentrations (grey), for each month with 95% confidence intervals, where the standard error is calculated over all measurements available for that hour of the day during that particular month.

1.6 Fossil Fuel and NEE Contributions

The contributions of the fossil fuel and NEE fluxes to the modelled concentrations are provided for Robben Island and Hangklip for the months July 2012 to June 2013. As for March 2012 to June 2012, the posterior estimates of the NEE fluxes are increased in such a way that the uptake of CO_2 cancels out the emission of CO_2 due to fossil fuel sources. Little adjustment is made to the contribution from the fossil fuel fluxes to the modelled concentration by the inversion.


Prior flux estimates Robben Island
July 2012
Posterior flux estimates Robben Island
July 2012
August 2012
September 2012
November 2012
February 2013
CO2 Concentration
Hour

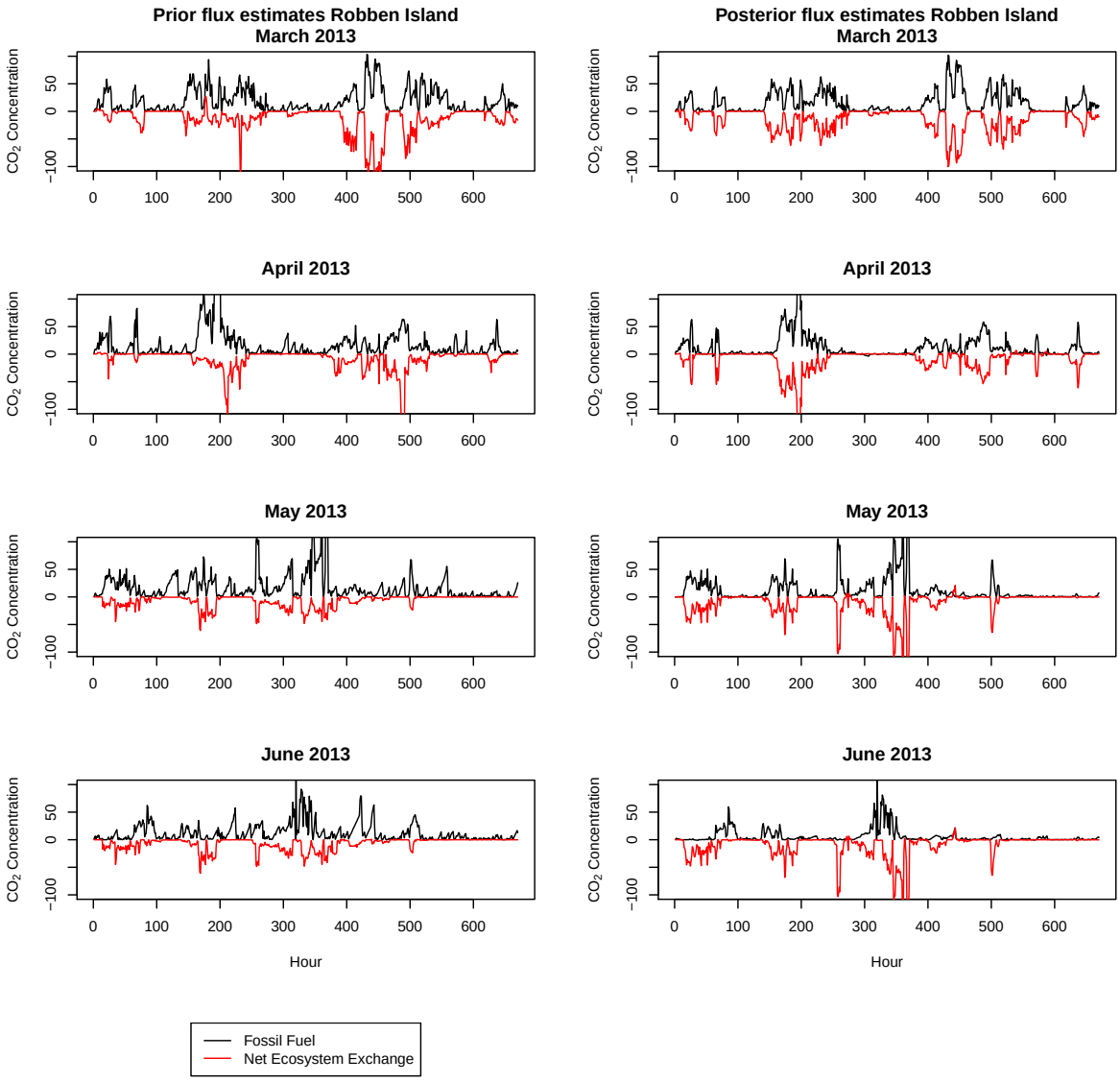
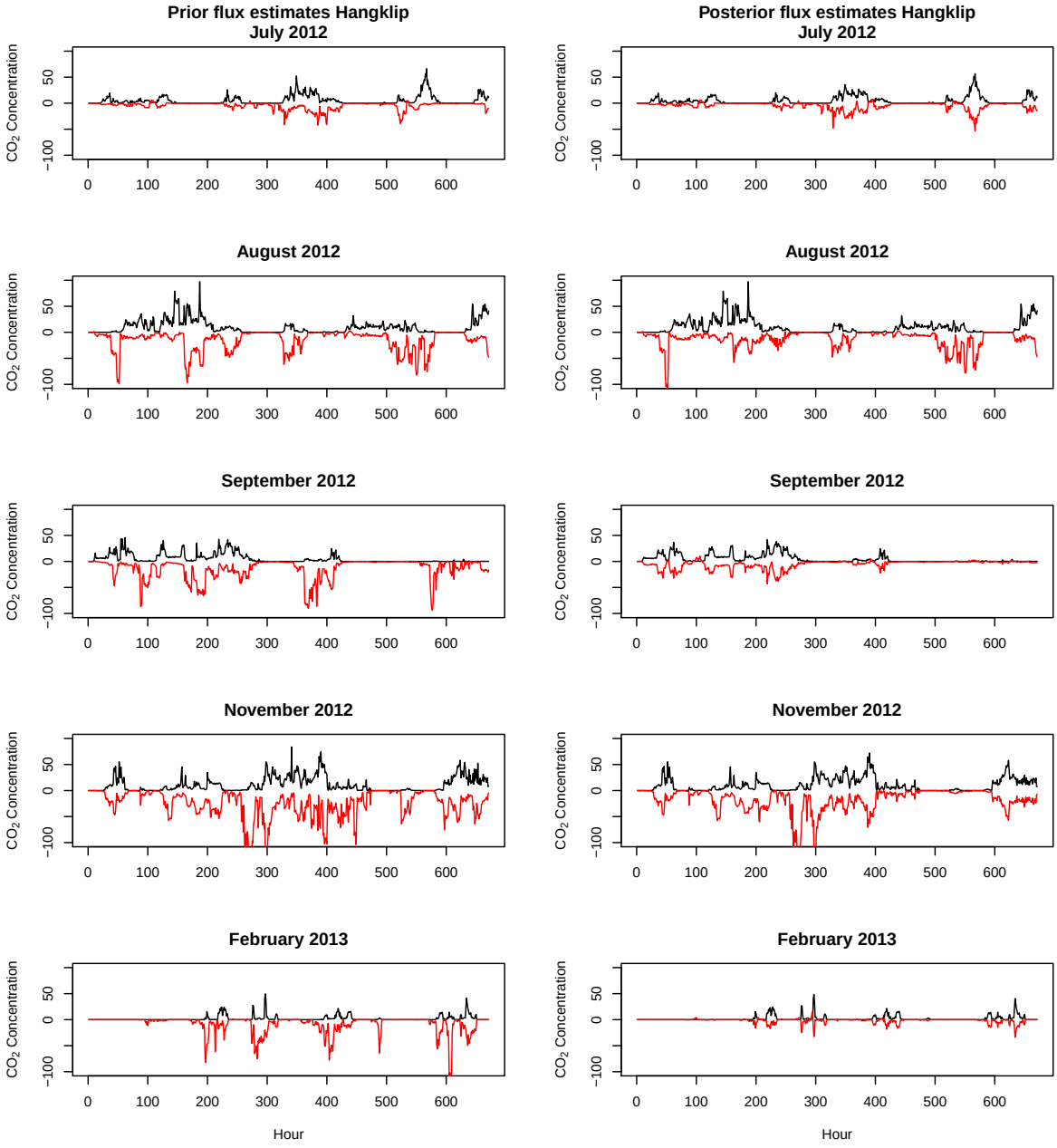


Figure 9. Contribution of the fossil fuel and NEE surface fluxes to the modelled CO_2 concentrations (ppm) at Robben Island.


Prior flux estimates Hangklip
July 2012
Posterior flux estimates Hangklip
July 2012
August 2012
August 2012
September 2012
September 2012
November 2012
November 2012
February 2013
February 2013
CO2 Concentration
Hour
−100
0
50
0
100
200
300
400
500
600

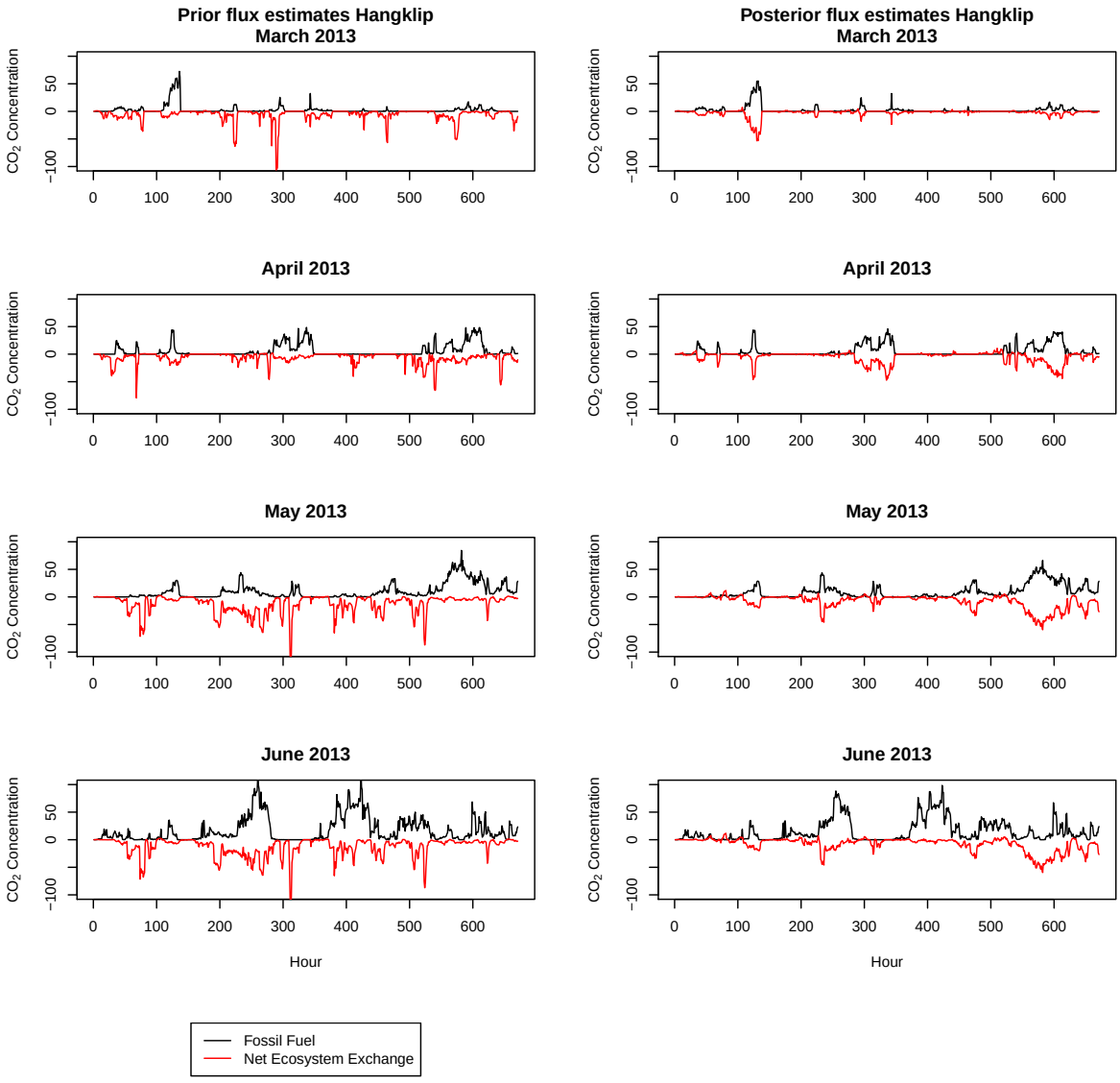


Figure 10. Contribution of the fossil fuel and NEE surface fluxes to the modelled CO_2 concentrations (ppm) at Hangklip.

1.7 Weekly Flux Estimates

Section 3.2 provides a summary of the estimates presented here for the pixel-level weekly fluxes over the full domain for each month.

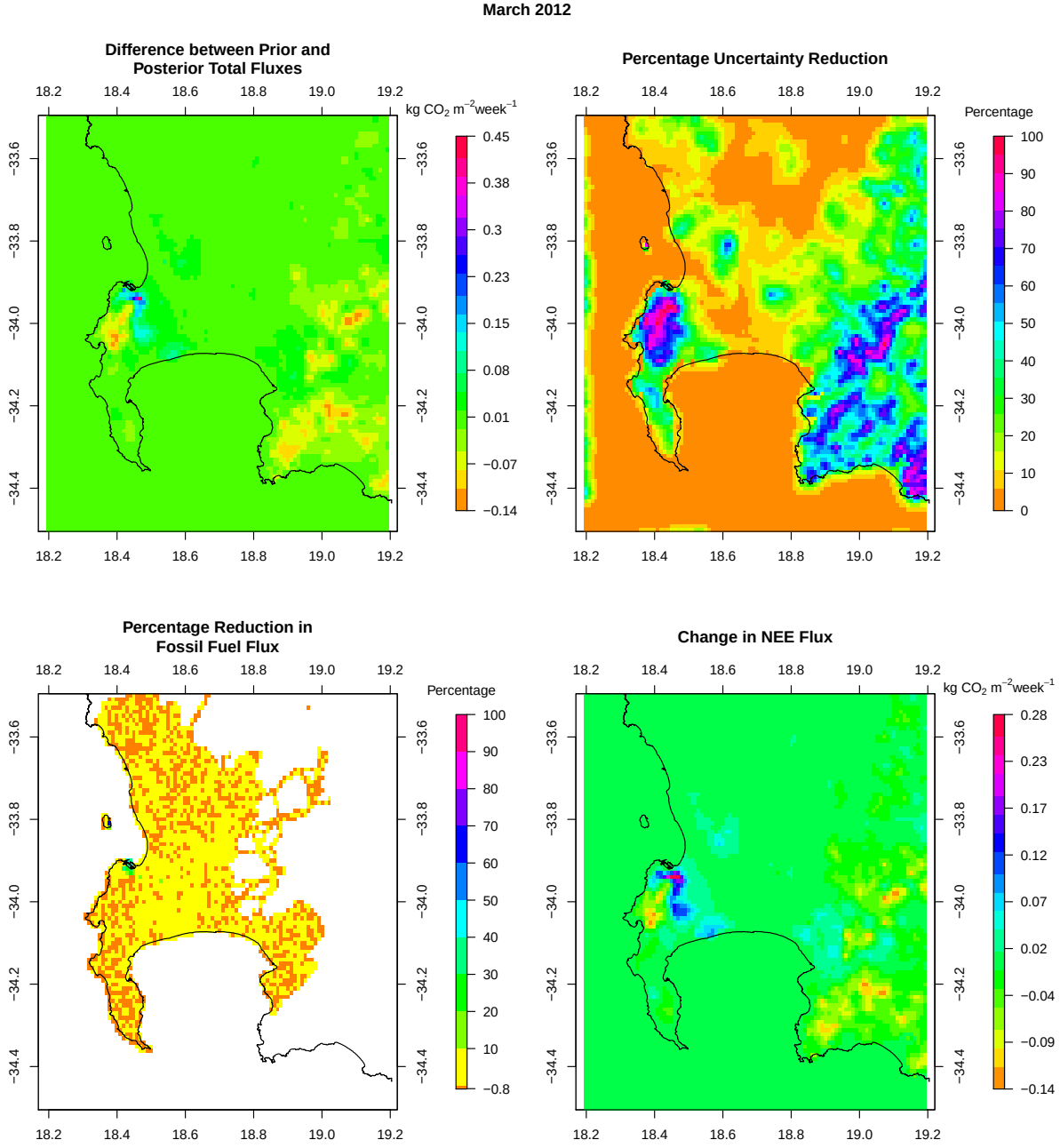


Figure 11. (Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for March 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

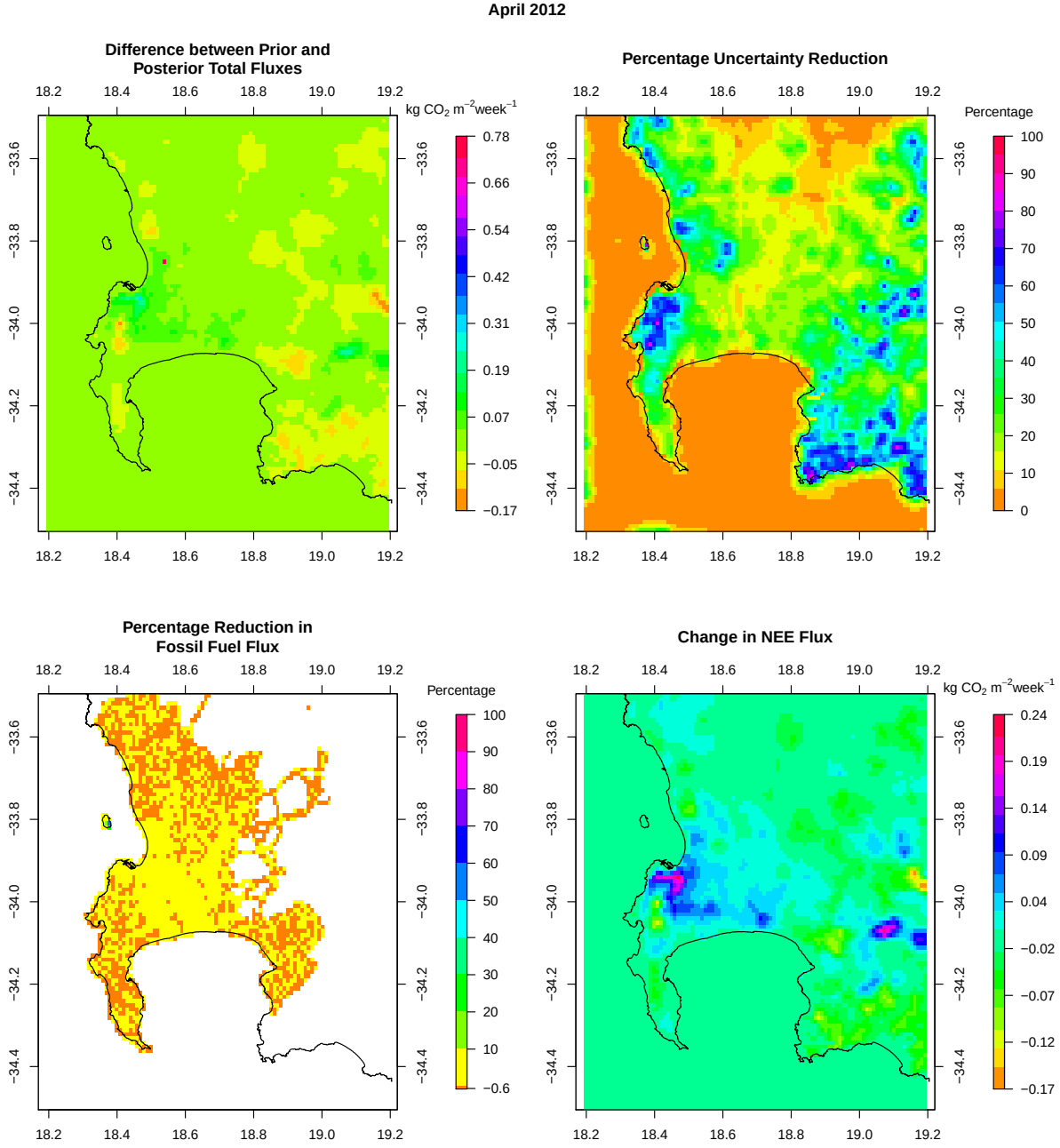


Figure 12. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for April 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

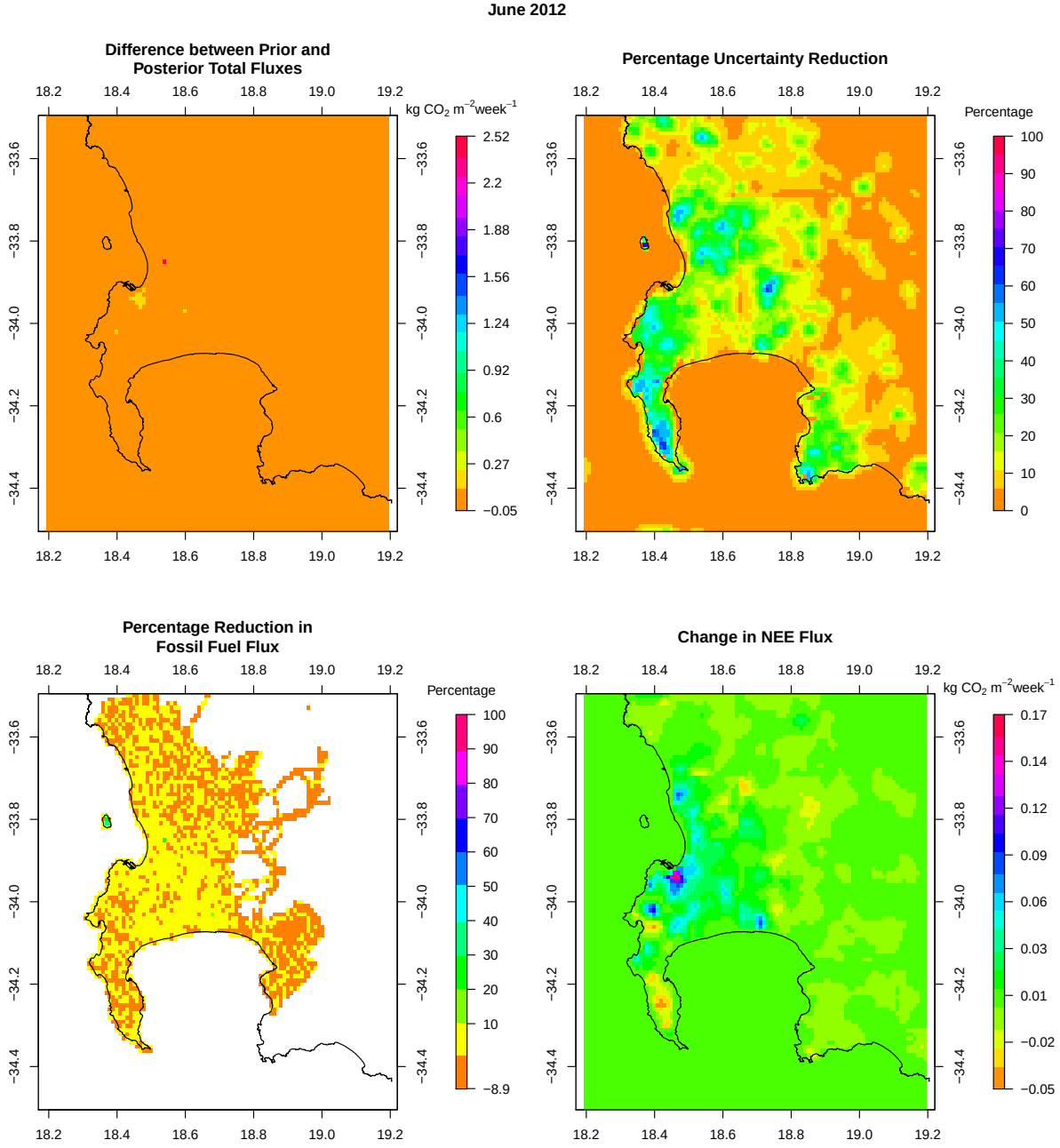


Figure 13. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for June 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior $\mathrm{CO_2}$ fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

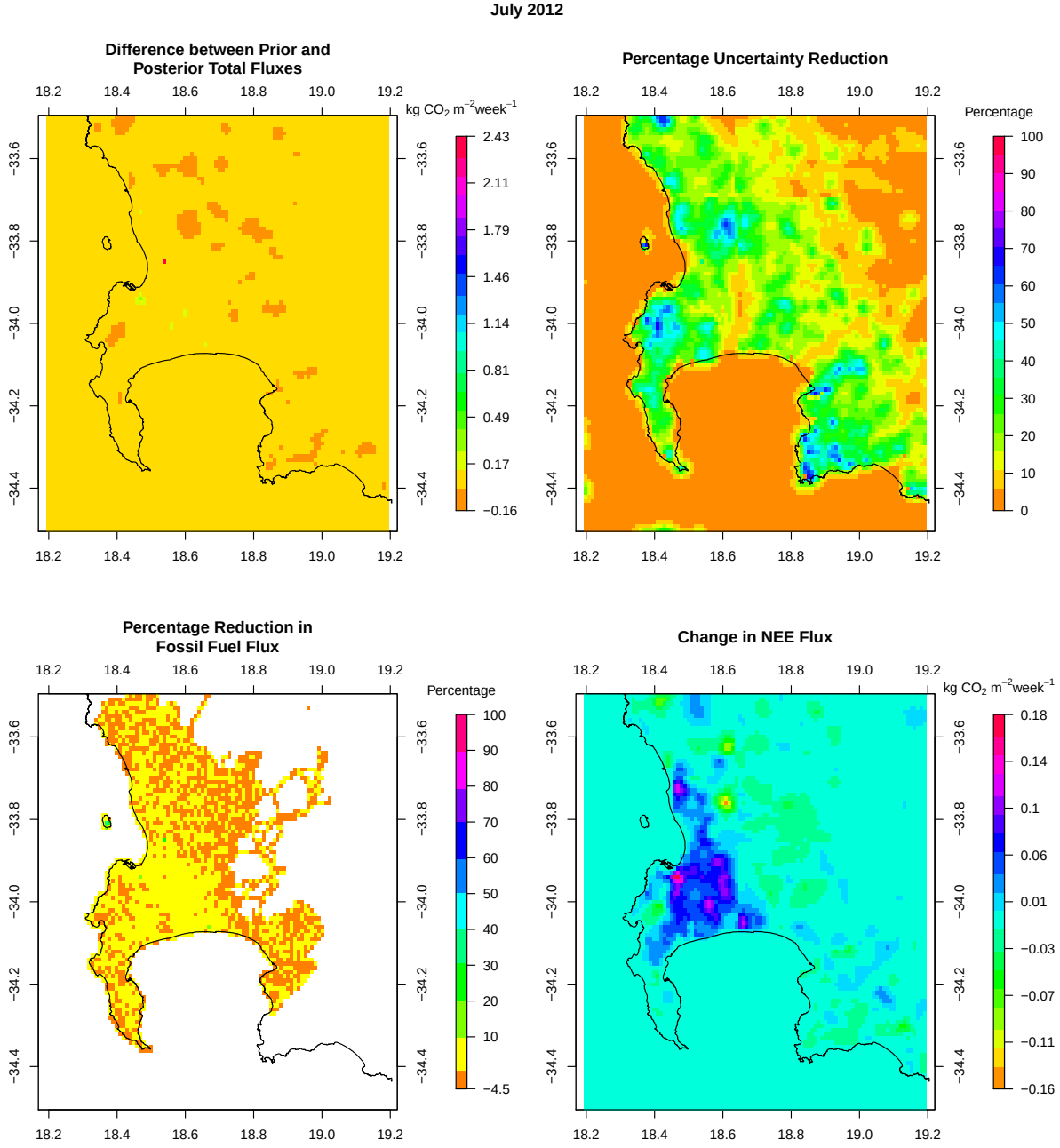


Figure 14. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for July 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

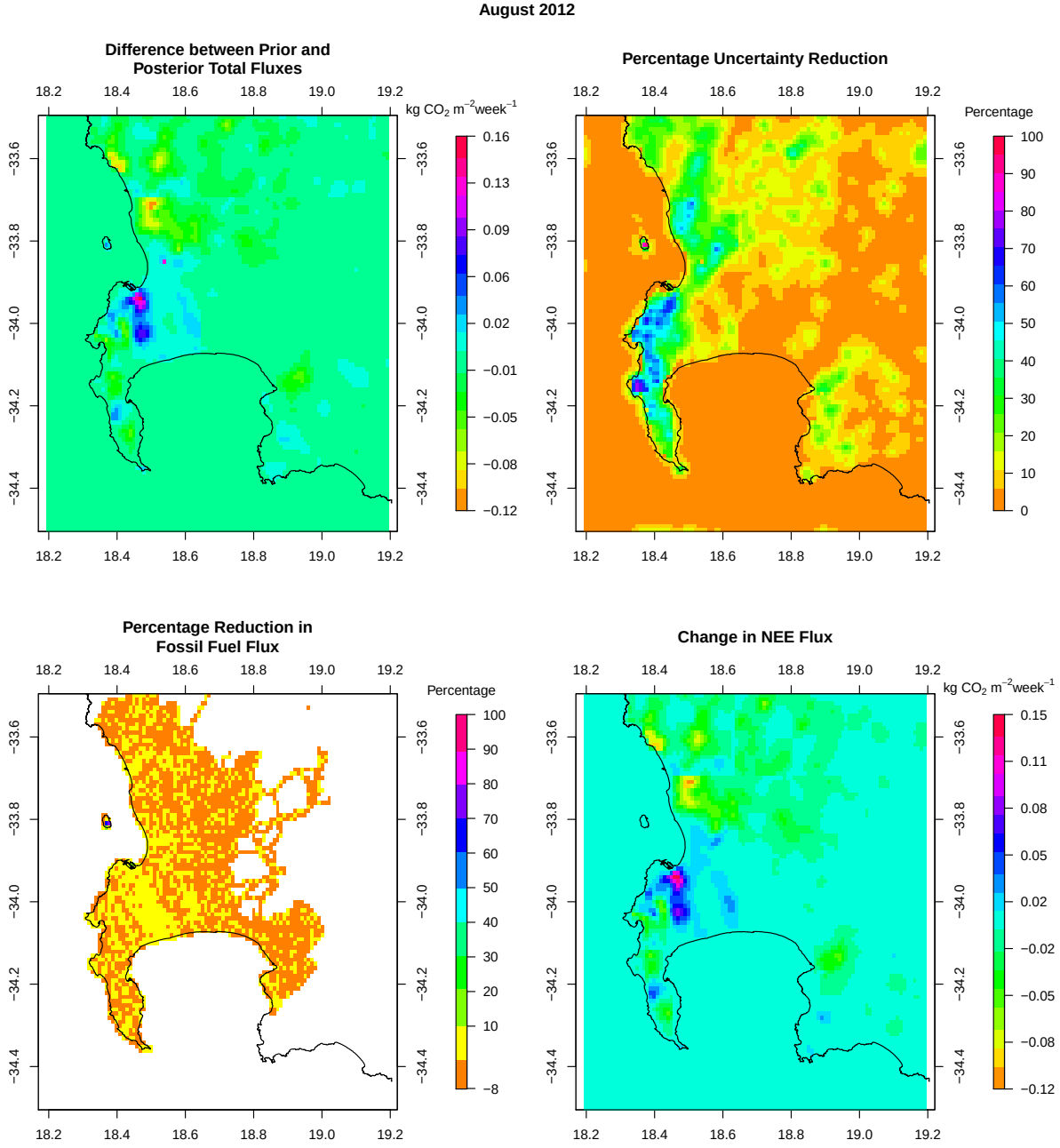


Figure 15. (Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for August 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

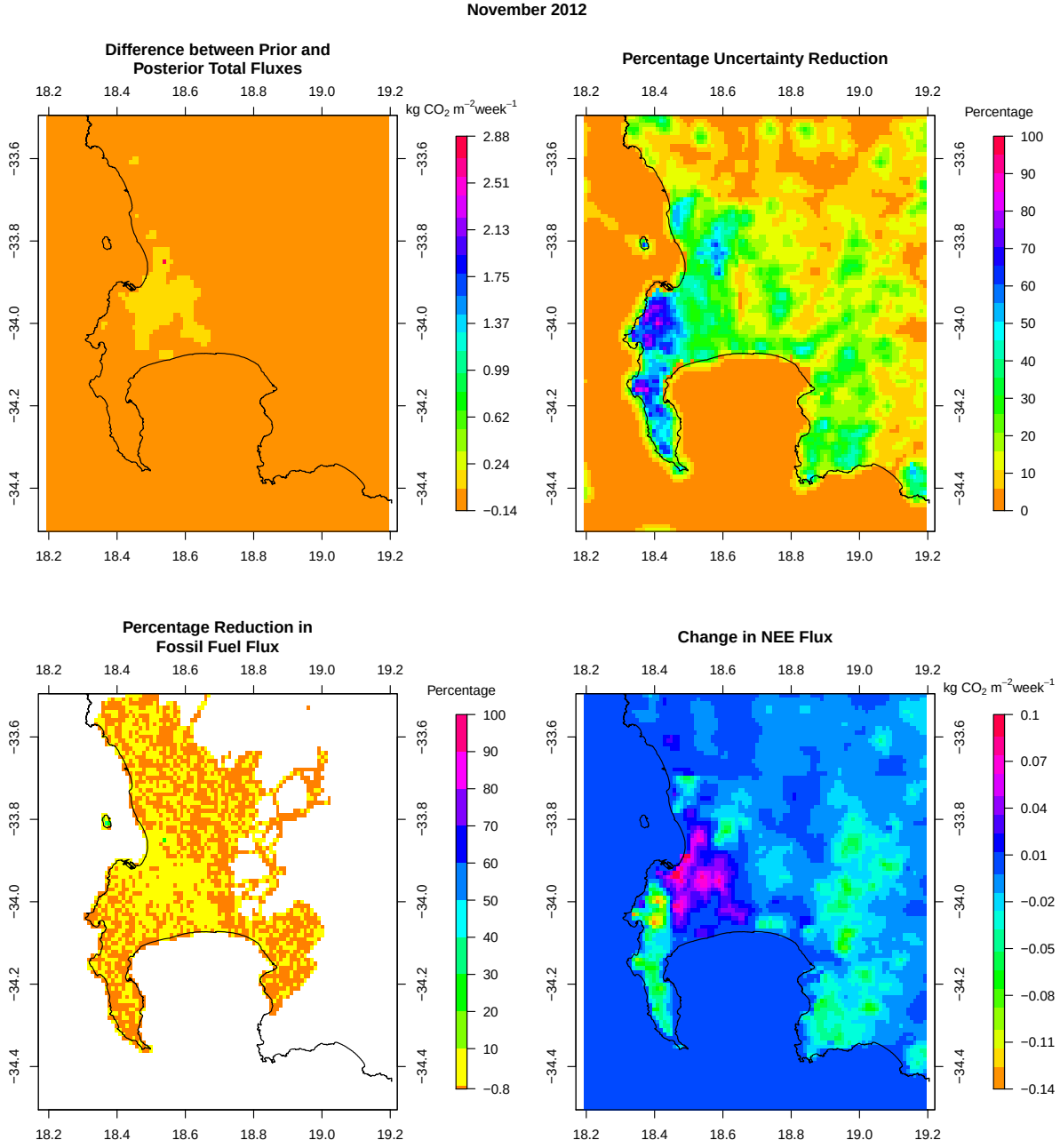


Figure 16. (Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for November 2012 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

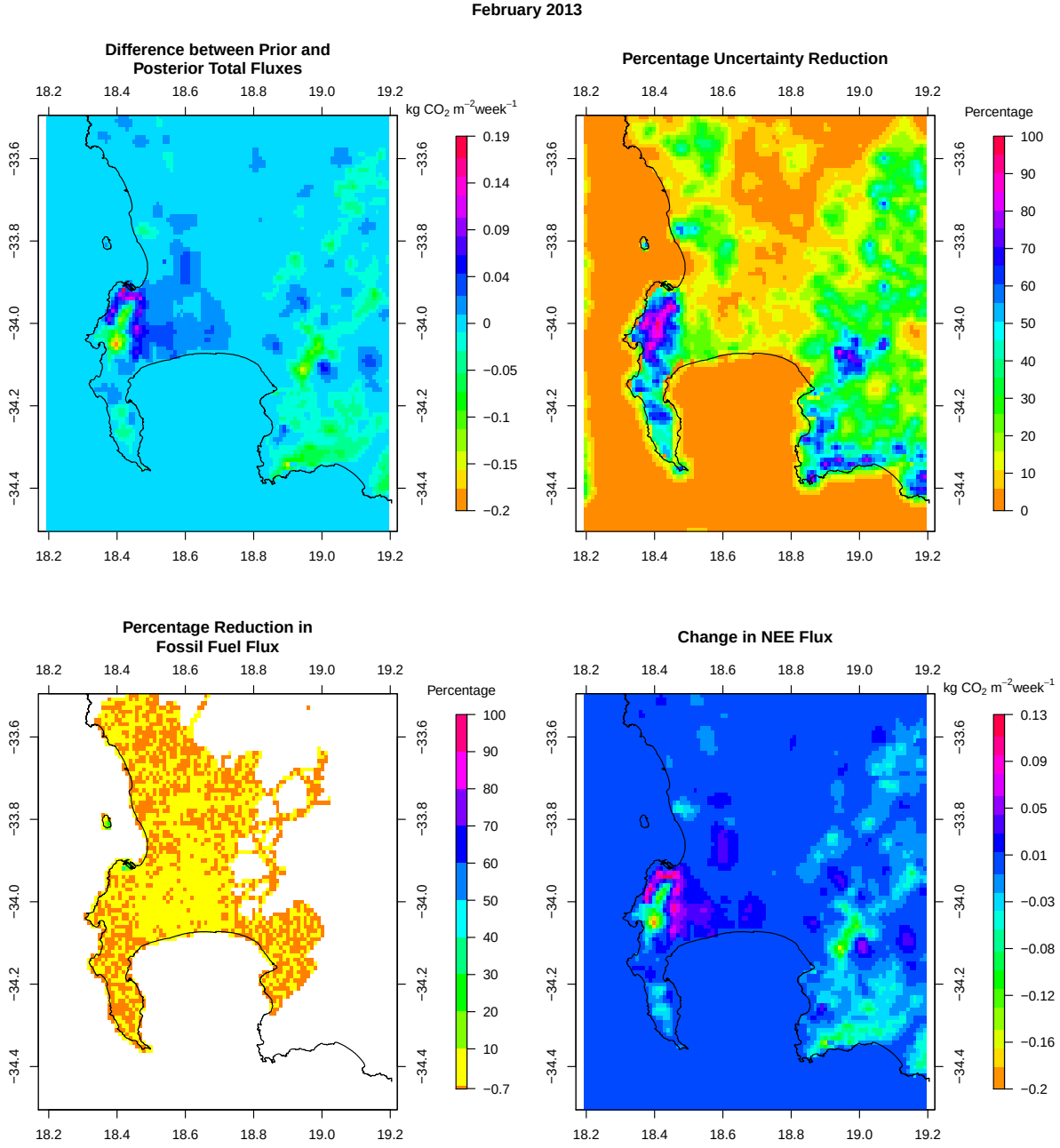


Figure 17. (Top left) Differences between the prior and posterior total flux estimates (kg $CO_2 m^{-2}$ $week^{-1}$) for February 2013 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) (kg $CO_2 m^{-2}$ $week^{-1}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

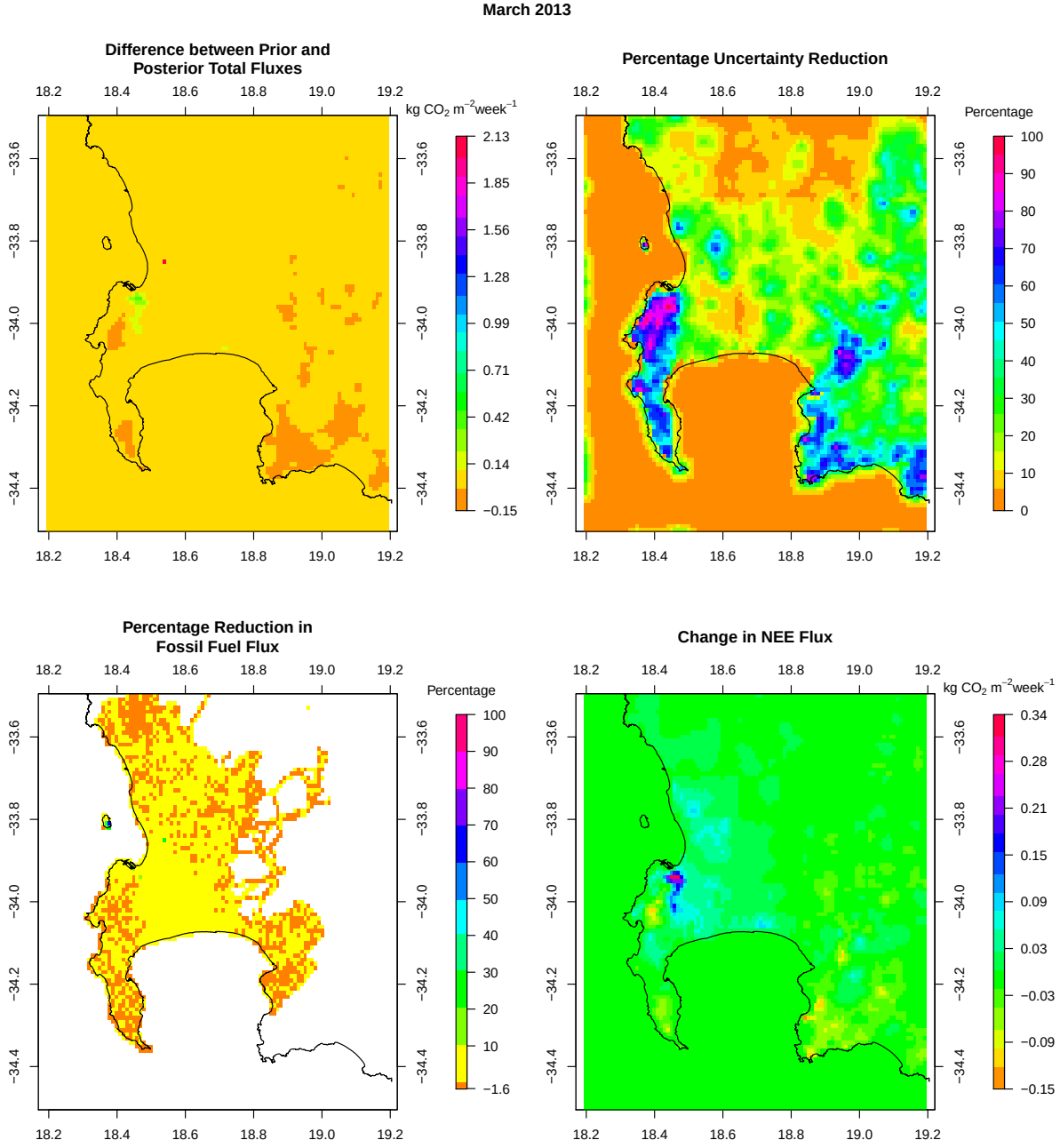


Figure 18. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for March 2013 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

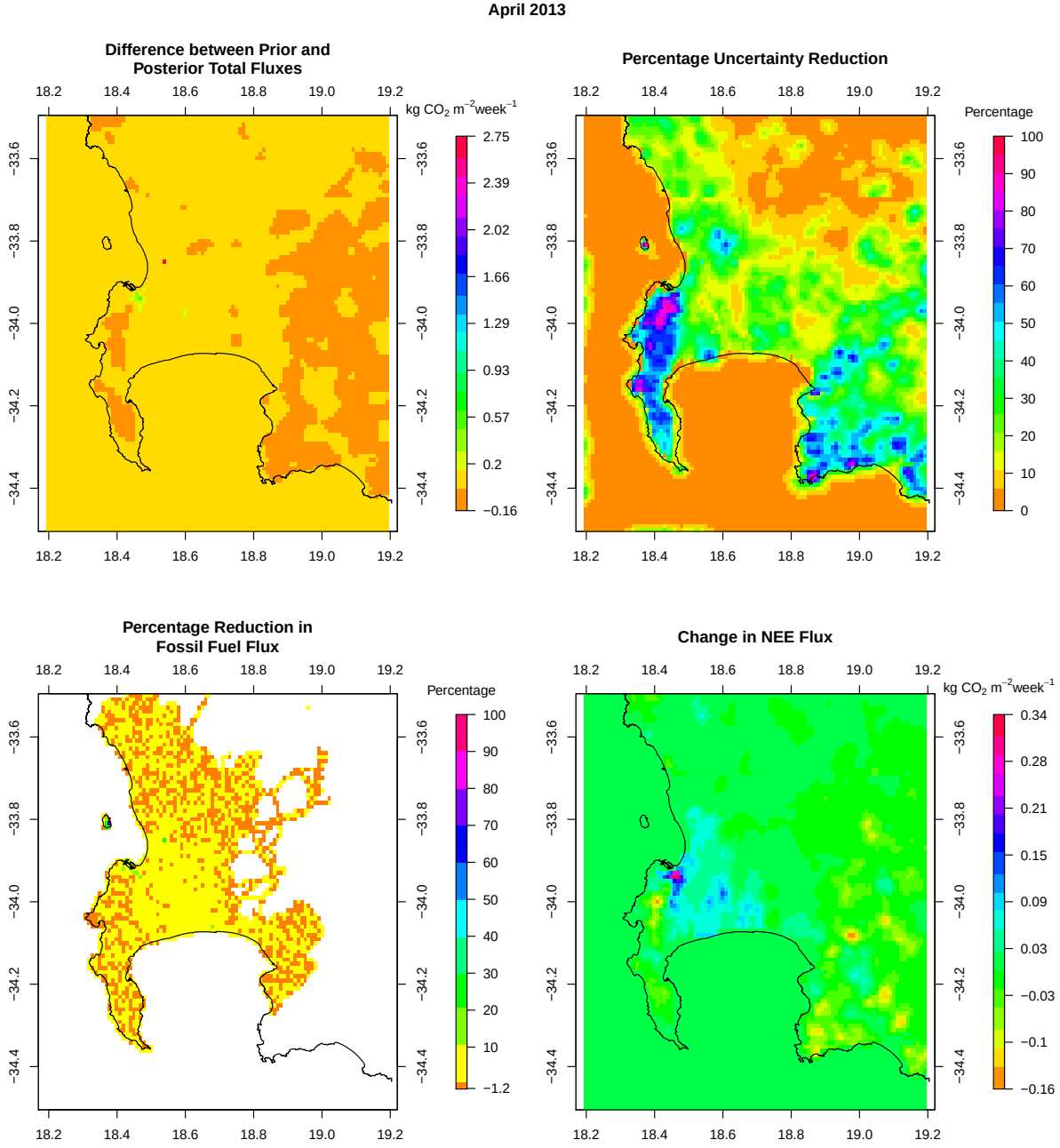


Figure 19. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for April 2013 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

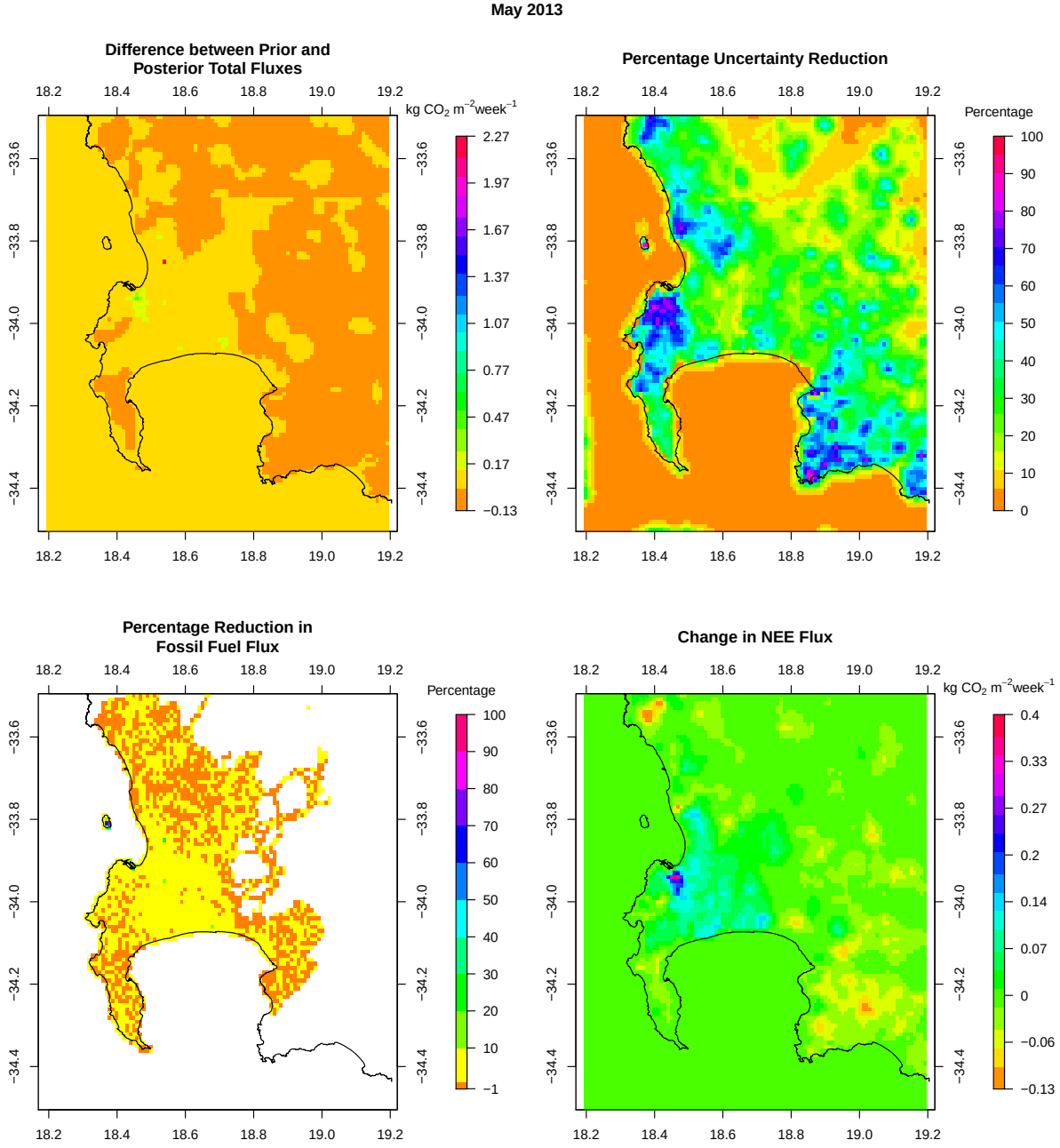


Figure 20. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for May 2013 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

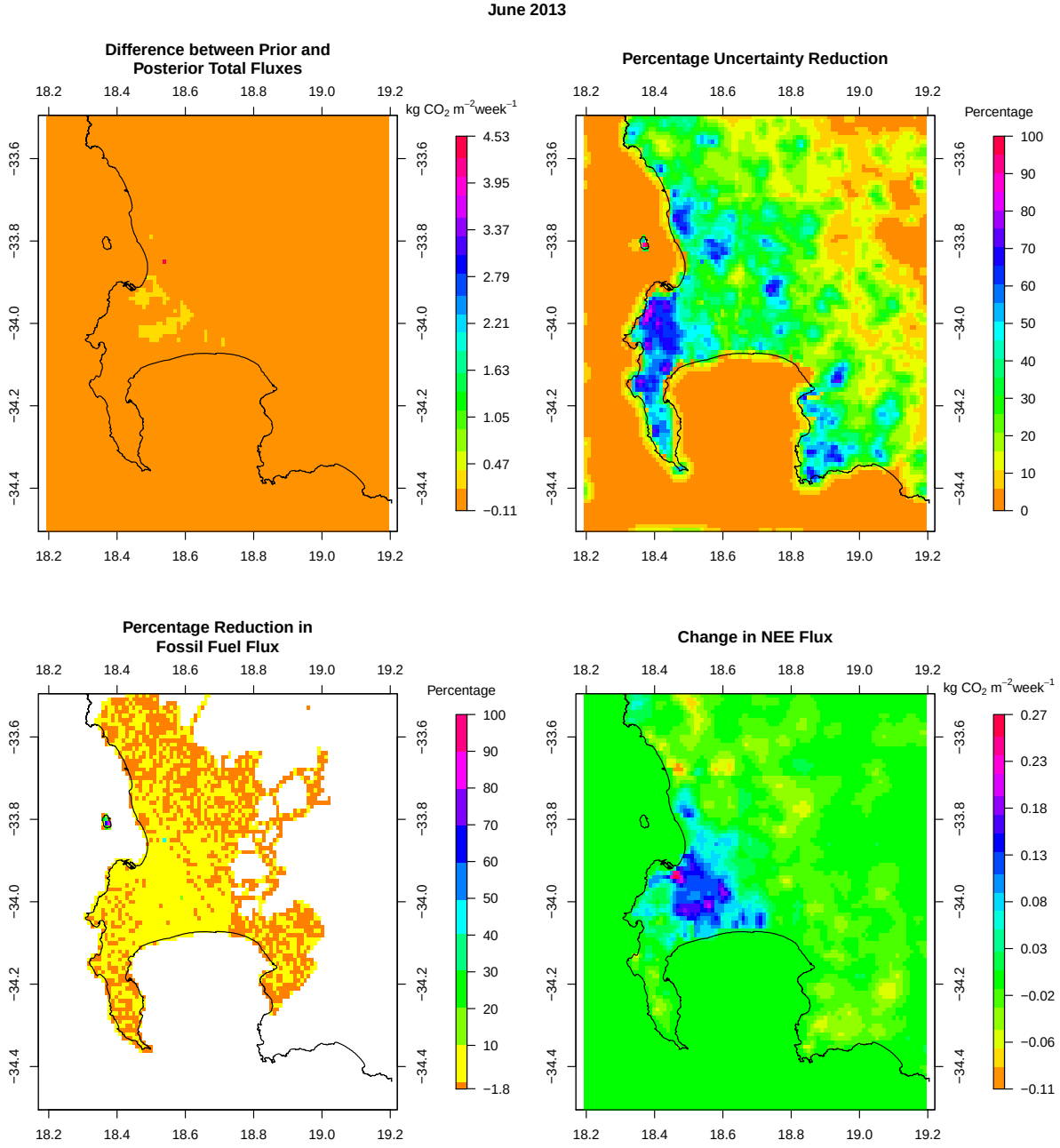


Figure 21. (Top left) Differences between the prior and posterior total flux estimates ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$) for June 2013 (prior - posterior). (Top right) Percentage reduction in the standard deviation of the flux estimate from prior to posterior. (Bottom left) Percentage reduction in the fossil fuel flux estimates from prior to posterior. (Bottom right) Differences in the biogenic flux estimates between prior and posterior estimates (prior - posterior) ($\mathrm{kg\ CO_2 m^{-2}\ week^{-1}}$), with negative values indicating posterior CO_2 fluxes were make more positive by the inversion compared with the prior estimates. Extent: between 34.5° and 33.5° south and between 18.2° and 19.2° east.

1.8 Toy Inversion

Let us consider an hourly measurement at a single site, with a fossil fuel flux daytime source, a fossil fuel flux night-time source, an NEE flux from the same location, and an NEE flux from a neighbouring pixel. We wish to solve for these four fluxes and the covariance matrix of the uncertainties in these fluxes. Selecting some of the most extreme values for the uncertainties and for the sensitivities for the current inversion framework we could get the following:

$$\mathbf{H} = \quad (0.0, 0.0126, 0.00902, 0.0032); \quad \mathbf{C_{s_0}} = \begin{pmatrix} 233 & 0 & 0 & 0 \\ 0 & 78 & 0 & 0 \\ 0 & 0 & 1578 & 1220 \\ 0 & 0 & 1220 & 1578 \end{pmatrix}; \quad \mathbf{C_c} = 4$$

Solving for the posterior covariance matrix of the flux uncertainties using:

$$\mathbf{C_s} = \left(\mathbf{H}^T {\mathbf{C_c}}^{-1} \mathbf{H} + {\mathbf{C_{s_0}}}^{-1}\right)^{-1} \tag{1}$$

$$= \mathbf{C_{s_0}} - \mathbf{C_{s_0}} \mathbf{H}^T \left(\mathbf{H} \mathbf{C_{s_0}} \mathbf{H}^T + \mathbf{C_c}\right)^{-1} \mathbf{H} \mathbf{C_{s_0}}. \tag{2}$$

gives

$$\mathbf{C_s} = \begin{pmatrix} 233 & 0 & 0 & 0 \\ 0 & 77.8 & -4.2 & -3.7 \\ 0 & -4.2 & 1500.2 & 1151.1 \\ 0 & -3.7 & 1151.1 & 1517.0 \end{pmatrix}; \quad \rho_{\mathbf{matrix}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -0.01 & -0.01 \\ 0 & -0.01 & 1 & 0.76 \\ 0 & -0.01 & 0.76 & 1 \end{pmatrix};$$

Although the sensitivity of the concentration measurement to the fossil fuel and NEE fluxes are not that different, the posterior covariances are small because the transport Jacobian only projects fluxes from individual pixels weakly into modelled concentrations. The uncertainties in the prior modelled concentrations that are attributed to the flux contributions ($\mathbf{H} \mathbf{C_{s_0}} \mathbf{H}^T$) are small relative to the uncertainties specified for the modelled concentration errors ($\mathbf{C_c}$). If we reduce the elements of $\mathbf{C_c}$ then the posterior covariances increase. For example, If $\mathbf{C_c}$ = 1 then

$$\mathbf{C_s} = \begin{pmatrix} 233 & 0 & 0 & 0 \\ 0 & 77.2 & -14.5 & -12.9 \\ 0 & -14.5 & 1310.0 & 982.8 \\ 0 & -12.9 & 982.8 & 1368.0 \end{pmatrix}; \quad \rho_{\mathbf{matrix}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -0.04 & -0.04 \\ 0 & -0.04 & 1 & 0.73 \\ 0 & -0.04 & 0.73 & 1 \end{pmatrix};$$

References

Lamigueiro, O., P.: Visualization methods for raster data, ver. 0.41. http://oscarperpinan.github.io/rastervis